CATALYSIS AND INHIBITION OF DNA POLYMERASE ETA, A TRANSLESION POLYMERASE IMPORTANT FOR CHEMOTHERAPEUTIC EFFICACY

By
Mark Thomas Gregory

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

October 2014

Preface and Acknowledgements

I feel extremely grateful that for doctoral thesis research I have been able to work in a laboratory with highly skilled and supportive people. I would like to thank the current and past members of the Yang group for the assistance, support and friendship they have given me during my tenure with them – including Drs. Lan Tian, Young-Sam Lee, Filip Golebiowski, Jinjun Wu, Min-Sung Kim, Samarendra Singh, Ye Zhao, Santiago Ramón-Maiques, Teruya Nakamura, Mikalai Lapkouski, Shikha Gupta, Joyce Cheung, and Christian Biertümpfel. I would also like to recognize the Laboratory of Molecular Biology at the NIDDK for fostering a stimulating research environment in which it has been my pleasure to work for the past 4 and a half years.

I would like to specifically thank my advisor Dr. Wei Yang. Her direction, encouragement, and guidance throughout my thesis work were instrumental in my success. I will be forever grateful.

The JHU-GPP directors Michael Lichten and Orna Cohen-fix and the NIH GPP offices in the OITE provided me with extraordinary support, particularly in the early stages of my thesis work. Their advice and assistance with administrative, personal, and scientific aspects of graduate life were critical as my thesis research progressed and evolved. Without their efforts this work would not have been possible.

I would also like to thank my graduate thesis committee, including Drs. Sarah Woodson and Joel Schildbach as well as past members Drs. Fred Dyda, Trina Schroer, Carole Bewley and James Hurley for their advice throughout my thesis and for spending time to read this dissertation. The guidance I received from these committee meetings was invaluable, focusing my research and adding tremendous value to my graduate training.

Finally, I would like to dedicate this work to my wife, Shawna. Shawna has supported me during my entire graduate career, through the exciting breakthroughs and the difficult setbacks. Her patience, understanding, and belief in me have been incredible, particularly during the challenges and failures that have occurred along the way.

Thesis Committee

Dr. Wei Yang	Advisor, Laboratory of Molecular Biology, NIDDK/NIH
Dr. Sarah Woodson	CMDB/Biophysics/JHU
Dr. Joel Schildbach	CMDB/Biology/JHU
Dr. Michael Lichten	Microbial Genetics and Biochemistry Section, NCI/NIH

Thesis Readers

Drs. Wei Yang and Sarah Woodson

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. DNA Polymerases and their role in replication

DNA replication is the process by which genetic material and information is copied. Replication underpins the evolutionary process and is conserved in all life yet discovered. At its simplest, DNA replication is the copying of a template DNA by synthesis of a new, complimentary DNA strand[1]. This process is carried out by a class of enzymes known as polymerases[2].

Polymerases were first discovered by Author Kornberg in 1955 when he isolated DNA Polymerase I (Pol I) from E. coli based on the characteristic ability to polymerize individual nucleotides[3]. Studies on polymerase isolates determined that dNTP's were the building blocks required by polymerases, not dNMP's as were originally thought based on the structure of DNA solved by Watson and Crick only a few years earlier[1, 3, 4]. Once the polymerase reaction was reconstituted in vitro, both Mg^{2+} and template DNA were found to be essential cofactors[3]. The requirement for DNA as a cofactor reinforced the Watson-Crick and Meselson-Stahl models for replication, using a template strand as a guide[5, 6]. Finally, the directionality of synthesis was determined to be 5´ to 3´ only[5].

These core characteristics, defined for the first discovered polymerase, Pol I, have remained true for all known DNA polymerases, even as the number of known polymerases has expanded[7]. Multiple polymerases have been discovered within single organisms as well as in each new organism discovered. For example, five DNA polymerases have been discovered in E. coli and so far,

and 17 in humans[8, 9]. In humans, these DNA polymerases are not redundant but serve unique roles[2].

The DNA polymerase super-family consists of 7 sub-families: A, B, C, X, Y, RT, and the recently described PrimPol family. Most polymerases of the A-, B-, and C- families have high fidelity and high processivity and are responsible for replication of chromosomal DNA[2]. In humans, these families include all the polymerases necessary for replication of undamaged DNA: Pol α, δ, ε for the genomic DNA and Pol γ, the mitochondrial DNA polymerase[10]. The X-family polymerases generally function in DNA repair pathways such as base excision repair (BER) in the case of Pol β, as well as non-homologous end joining (NHEJ) in the cases of Pol λ and μ[11]. These polymerases resynthesize short stretches of DNA removed during repair[11]. The Y-family polymerases are also important for DNA repair. Y-family polymerases are known as translesion synthesis (TLS) polymerases because they bypass DNA lesions that stall replication forks, due to the inability of the high fidelity replicative polymerases to synthesize opposite the lesions[12]. The RT sub-family, which includes Reverse Transcriptase from viruses as well as telomerase in eukaryotes, functions to synthesize DNA from RNA templates[13, 14]. The final sub-family is the PrimPol family, characterized by having both primase and polymerase activity[9, 15]. All polymerases are categorized into sub-families based on sequence homology. Each sub-family is defined by a high degree of sequence similarity among its members, but homology between sub-families is exceptionally low given their shared chemistry[16]. In fact, similarity between polymerase sub-families was too low for

useful sequence guided discovery of the more recently found polymerase sub-families: X, Y, and PrimPol. The A, B, Y, and RT families share a similar fold while the X and C family polymerases appear to have evolved separately, having distinct folds[7, 15]. However, despite the low sequence homology and unrelated tertiary structures, a highly similar domain organization is seen across the polymerase families[7, 15].

The polymerase domain organization consists of three main subdomains; the palm, thumb, and finger [7] (Fig. 1). The protein folds of these subdomains vary between families, with the A, B, Y, and RT families sharing a similar fold, while the X and PrimPol families both have unique folds. Despite the different tertiary structures, the active site composition and subdomain configuration are highly conserved among all DNA polymerases, likely because the same chemical reaction is catalyzed by these enzymes (Fig. 1). The thumb subdomain is involved in DNA binding, associates with the upstream DNA duplex, and contributes to processivity[17]. This subdomain is present in all polymerase families except the PrimPols, which consequently have very low processivity[15]. The thumb subdomain is adjacent to the palm subdomain, which houses the three conserved active site carboxylates. These acidic residues coordinate two Mg^{2+} ion cofactors in the active site that are required for polymerase catalysis. The active site is located in the palm subdomain at the interface of the palm and finger subdomains[7]. The finger subdomain is located directly above the nascent base pair of the templating base and incoming dNTP[7]. In high fidelity A-family polymerases, its positioning over the templating base allows the finger

subdomain to sterically sense the conformation of the nascent base pair and to determine if the correct dNTP is bound[18, 19]. This sensing is performed by an α-helix termed the O-helix, which sterically prevents full closure of the finger subdomain onto mismatched template:dNTP base pairs, contributing to the extremely high fidelity of replication[19]. Fidelity is further increased for some polymerases of the A, B, and C families by the presence of a 3´ to 5´ exonuclease activity, located either in a fourth structural domain or an associated subunit[20, 21], which removes incorrectly incorporated bases during replication (Fig. 1).

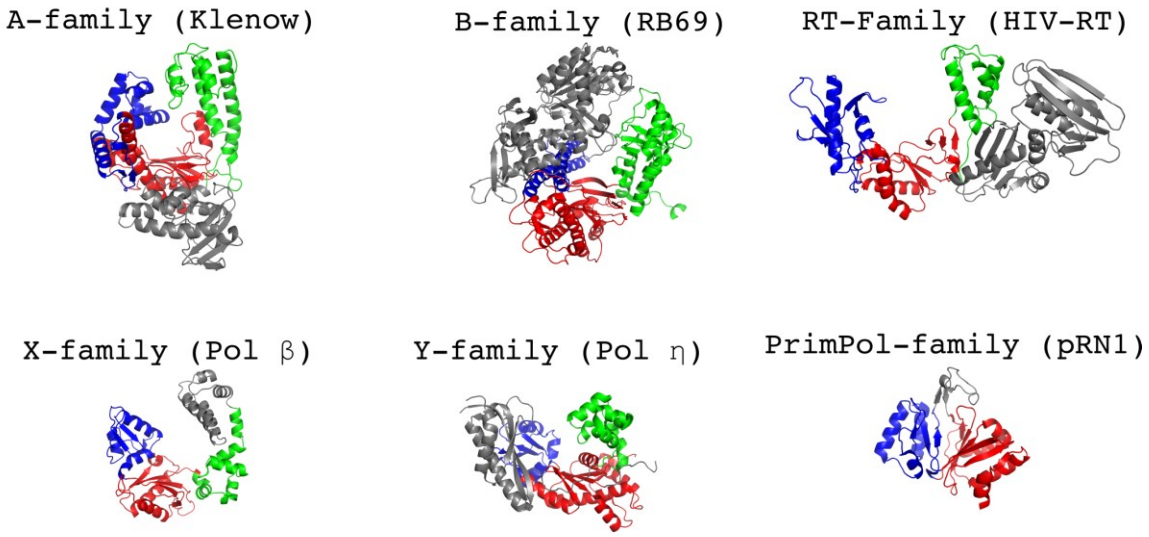


Figure 1. Conserved architecture of polymerases

The three polymerase domains are highlighted in color: Thumb in green, palm in red, and finger in blue. Non-conserved elements are shown in grey, including the proofreading exonucleases of the A and B families. Polymerases shown are Taq Klenow fragment (A-family, pdb code 3KTQ[22]), RB69 (B-family, pdb code 4FK4[23]), Pol β (X-family, pdb code 4PPX[24]), Pol η (Y-family, pdb code 3MR2[25]), HIV-RT (RT-family, pdb code 1RTI[26]) and the Polymerase/Primase domain of pRN1 PrimPol (PrimPol-family, pdb code 1RNI[15]).

1.2. Polymerase active site and nucleotidyl transfer reaction

The similar domain organization of polymerases reflects the structural characteristics necessary for their shared chemistry, incorporating free nucleotides (dNTP) into a growing DNA primer strand. The protein fold of the palm domain varies between polymerase families (Fig. 1), however despite different folds, the active sites of all DNA polymerases that have been structurally characterized share a similar active site composition [15, 25, 27, 28]. The conserved active site core is composed of three carboxylates arranged in a triangle like pattern to coordinate two Mg^{2+} ion cofactors [15, 28]. The positive Mg^{2+} ions in turn bind the highly negative tri-phosphates of the incoming dNTP in polymerase-DNA-dNTP ternary complex crystal structures (Fig. 2)[29].

Recently, using time resolved crystallography, the dynamics of polymerase catalysis was observed using Pol η by Nakamura et al (Fig. 2)[27]. They constructed a pre-reaction complex consisting of the polymerase complex with DNA, incoming nucleotide, and only one metal bound in the B-site. For experimental reasons the metal in their structure was a Ca^{2+} ion instead of the Mg^{2+} ion which would be present *in vivo*. The metal bound at the B-site coordinates the α, β, and γ phosphate oxygens[25]. This ion is part of the Mg^{2+}dATP complex in solution and thus is the first metal ion to bind with dNTP in the active site[30, 31]. They termed this complex the Ground State (GS). In the GS, both the dNTP and primer are bound but not aligned for catalysis, specifically the primer 3´-OH is too far from the α-phosphate, at over 4Å, to react[27]. The second Mg^{2+} ion, sometimes termed the catalytic Mg^{2+}, next binds at the A-site to

coordinate the α-phosphate and the 3´-OH from the terminal base of the primer strand[25, 29]. This state, with two Mg^{2+} ions bound, was termed the Reactant State (RS). In the RS, the primer 3´-OH has shifted nearer the α-phosphate, to ~3.4Å, due to coordination with the A-site Mg^{2+} ion, and is aligned for nucleophilic attack (Fig. 2)[25, 27, 32]. Once the polymerase transitioned fully from the GS to RS, the primer was deprotonated and underwent nucleophilic attack on the α-phosphate, forming the nascent bond between the primer and incoming nucleotide and cleaving the α- to β-phosphate bond[27, 33]. This is believed to be an S_N2 type reaction, having a penta-covalent α-phosphate transition state intermediate[34, 35]. The transition state is stabilized by interaction with both Mg^{2+} ions, which directly coordinate one of the negatively charged, non-bridging oxygens of the α-phosphate[27, 36].

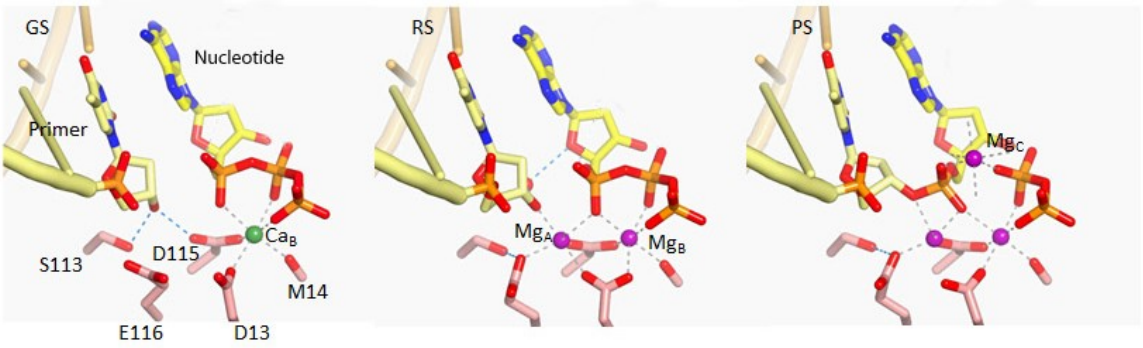


Figure 2. Polymerase Active Site and time resolved catalysis

Time resolved structures of WT polymerase η catalyzing the addition of dATP onto a DNA primer solved by Nakamura et al (pdb codes: 4ECQ, 4ECS, and 4ECV)[27]. Protein residues are colored in salmon, the primer and incoming dATP nucleotide are colored in yellow, and the calcium and magnesium metal ions are colored in green and purple, respectively. The Ground State (GS) is shown in the left panel prior to Mg^{2+} addition. The Reactant State (RS) is shown in the middle panel and the Product State (PS) is shown in the right panel. In the Product State a third Mg^{2+} ion was observed.

Pre-steady state kinetic analysis of polymerase catalysis has been performed on a variety of DNA polymerases including Dpo4[37], Pol β[38, 39], Pol η[40], Pol I[41], and Taq[42] with a goal of defining the rate-limiting step in polymerase catalysis. Based on early polymerase crystal structures, it was originally widely assumed that a large conformational change, closing the finger domain onto the dNTP bound palm domain prior to the reaction, was rate-limiting[43]. However,

there is no direct data to support finger domain movement as the rate-limiting step. More recent FRET and fluorescence based assays on Pol β, Klentaq1, and RB69 have measured the rate of finger domain movement as much faster than the rate of polymerization, proving it is not the rate-limiting step in catalysis[38, 42, 44]. Currently, it is believed that either the chemical step or a subtle active site conformational change immediately prior to, and possibly associated with, the chemical step is rate-limiting[41, 44-47]. Unfortunately, due to the complexity of the polymerase active site there are limited means to assess conformational changes within the polymerase active site, such as altered enzymes or substrates. Consequently, the precise nature of the rate-limiting step has not yet been determined for polymerase catalysis[46]. In lieu of precise experimental assignment of the rate-limiting step, quantum mechanical/molecular mechanics (QM/MM) simulations on polymerase β catalysis have indicated that deprotonation of the primer 3´-OH may be rate-limiting[45]. It should be noted, however, that in my work steady state kinetic conditions are used. Previous work on yeast Pol η have suggested that in steady state, single nucleotide insertion assays the disassociation of the polymerase from the DNA is rate-determining[40]. Chemical rates will be masked by this slow step in steady state conditions, unless enzyme or substrate perturbations significantly reduce these rates.

The base responsible for deprotonation of the 3´-OH is not known for any polymerase and remains a critical question within the field. Deprotonation by active site residues using a general-acid-base mechanism [48, 49] as well by solvent waters[27, 45, 50, 51] have both been proposed. QM/MM studies of 3´-OH

deprotonation on the Dpo4 and Pol β polymerase active sites have consistently favored solvent water for deprotonation[45, 50, 51]. The argument for solvent water deprotonation is further strengthened by a proposal by Tom Steitz that coordination of the 3´-OH to the Mg^{2+} ion, serving as a Lewis acid, lowers the 3´-OH's pKa dramatically, removing the requirement for a strong general base for deprotontation[28]. Mg^{2+} ions serve as Lewis acids in a variety of other enzymatic reactions, lowering the pKa of catalytic groups to enable chemistries to occur that would otherwise require non-biological pH extremes[52, 53]. The time-resolved crystallographic study of Pol η observed a water, which was termed the transient water, transiently associate with the 3´-OH and temporally correlate with product formation[27], suggesting this water may accept the 3´-OH. Nakamura and colleagues blocked binding of the transient water using an RNA terminated primer, which would occlude the transient water site with the 2´-OH, but did not observe significant activity loss[27]. However, they hypothesized that the O2´ may relay the proton to bulk solvent instead; suggesting this route for the proton is still viable and requires additional interrogation.

In a similar time-resolved experiment on the DNA polymerase Pol β, no transient water was seen during the early time points of the reaction, leading the authors to suggest one of the conserved aspartates deprotonate the 3´-OH[33]. However, the reaction occurred so rapidly no time points could be collected prior to product formation when deprotonation, and any transient water interactions, would occur. Additionally, investigation of any of the three conserved acidic

residues is impossible with traditional mutagenesis methods used to identify general acids or bases, due to their role coordinating the Mg^{2+} ions.

One other possible deprotonation candidate was identified in the Pol η time-resolved structures: S113[27]. S113 hydrogen bonds directly to the primer 3´-OH in the GS. The S113 position of Pol η is highly conserved across the polymerase superfamily as a serine in the Y-family, histidine in the A-family, and threonine in the B-family[54]. Mutation of this site in *E. coli* Pol I resulted in reduced, but not totally eliminated, activity *in vitro*, suggesting this residue is not serving as the general base in polymerase catalysis[55]. In Pol η, an S113A mutation was shown *in vitro* to have only minimal activity loss, much less than observed for the Pol I mutation at this site[27]. However, despite the high conservation of this site and the activity loss upon mutation no role has been described for this site.

In this work I combined mutational approaches with time-resolved crystallography of Pol η to directly observe the effect of site specific perturbations, of both protein residues and DNA nucleotides, on the polymerase catalysis dynamics. This study focused on conclusively identifying and validating the moiety that accepts the 3´-OH proton upon deprotonation.

1.3. The translesion synthesis (TLS) polymerase Pol η

Polymerase η is a member of the Y-family of DNA polymerases. Pol η was discovered in patients presenting with Xeroderma Pigmentosum (XP) syndrome[56, 57]. XP patients suffer from extreme sensitivity to any amount of

sunlight and increased incidence of keratoses and skin cancer that has been linked to deficient NER dependent repair of DNA damage, specifically UV light induced cyclobutane pyrimidine dimer (CPD) lesions[58]. A subset of XP patients, however, had perfectly functional NER, and instead lacked functional Pol η. Consequently, Pol η was termed the Xeroderma Pigmentosum Variant (XPV) gene[59, 60]. Pol η deficient XPV cells have reduced survival upon UV treatment, indicating that Pol η is involved at the replication fork in bypassing CPD lesions[58]. Pol η, as well the other Y-family DNA polymerases, are known as translesion synthesis (TLS) polymerases, due to the intrinsic ability to synthesize past lesions present on the template strand[61]. Each Y-family polymerase is specialized for a specific DNA lesion or subset of lesions[12]. *In vitro* characterization of Pol η has confirmed Pol η's role in bypass of CPD lesions[25, 56, 62].

For Pol η and other Y-family polymerases, access to the replication fork is a regulated event in the cell[63]. Pol η fidelity is orders of magnitude lower than that of the canonical replicative polymerases such as Pol ε, and if allowed to replicate large sections of the genome would result in mutations[64]. Instead, a mechanism termed polymerase switching is used when a replication fork encounters a lesion[65, 66]. UV exposure activates the Rad6/18 dependent ubiquitylation of the PCNA.[63]. Upon PCNA ubiquitination, Y-family polymerases are recruited to the fork[67]. Once the appropriate TLS polymerase binds, in this case Pol η, TLS past the CPD lesion rescues the stalled fork[25]. Y-family polymerases fortunately also have very low processivity, causing them to fall off

shortly after TLS of the lesion and to allow the high fidelity polymerase access to the fork[25].

The structure of yeast Pol η has been solved by the Aggarwal group and human Pol η by the Yang group [25, 32, 68]. Pol η contains all the conserved polymerase subdomains: thumb, palm, and finger (Fig. 3A). The 3´ to 5´ exonuclease domain seen in A- and B-family polymerases is not present in Pol η, which contributes Pol η's lower in fidelity[7, 21]. Instead Pol η, like all Y-family members, has a fourth subdomain termed the "little finger"[69]. The little finger subdomain functions to bind the template strand during DNA synthesis past CPDs and other bulky lesions[25]. The CPD lesion is large and distorts the B-form structure of the DNA double helix, causing misalignment of the active site [70]. However, within the little finger subdomain of Pol η, a beta sheet makes extensive contacts to the phosphate backbone of the template to reinforce B-form structure [25, 68]. Within the human Pol η structure, the DNA is bound on the opposite sides by the thumb and little finger subdomains, forming a molecular splint around the DNA that assists Pol η in positioning the lesion correctly in the active site[25]. The Pol η active site itself is also specialized for TLS of large adducts. Unlike the tight fitting active sites that ensure correct base pairing in the high fidelity polymerases, Pol η's active site is very open, which also contributes to its lower fidelity[25, 32]. The Pol η finger domain contains a large pocket above the templating base pair where adducts are accommodated during TLS[25, 32] (Fig. 3B). This pocket is unique to Pol η, and is what allows Pol η to accommodate CPD lesions, as well as other large lesions.

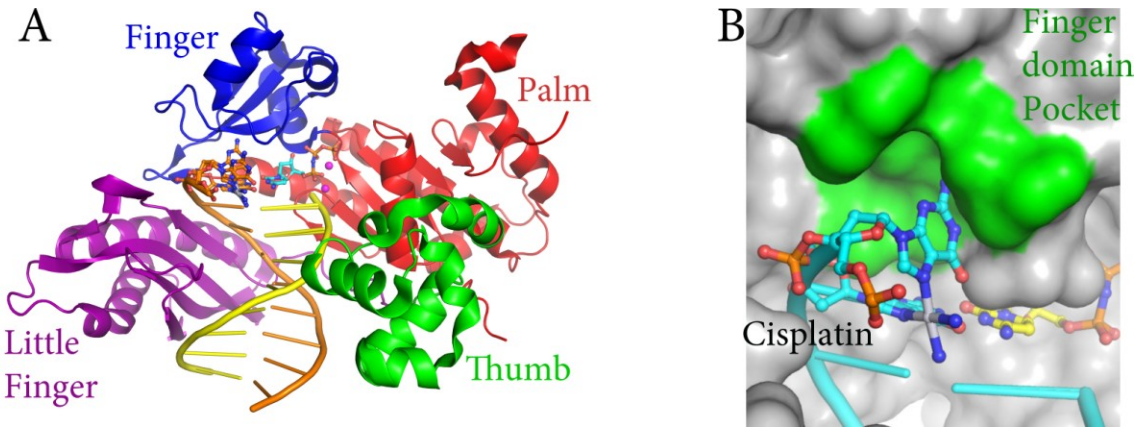


Figure 3. Polymerase η structure and bypass of a cisplatin lesion

This figure is adapted from Zhao et al structures, pdb code 4DL4 [32]. A. Structure of Pol η bound to cisplatin damaged DNA in the active site. The four domains of Y-family polymerases are highlighted with the thumb in green, palm in red, finger in blue and little finger in purple. The DNA is shown in yellow (primer) and orange (template) with the incoming nucleotide in cyan and damaged guanines in stick representation. B. The large active site of Pol η accommodates the cisplatin lesion during TLS. The CPD-binding pocket, unique to Pol η, is colored green. This pocket is also used in a similar manner to accommodate CPD lesions during TLS.

1.4. Platinum chemotherapeutics and cancer resistance to platinum compounds

The platinum drugs cisplatin, carboplatin, and oxaliplatin are used in the clinical treatment of approximately half of all cancer patients that receive chemotherapy [71]. These platinum chemotherapeutics function mainly by binding to and damaging genomic DNA, primarily forming bifunctional intrastrand crosslinks[72]. Platinum-DNA lesions cannot be bypassed by high fidelity replicative polymerases, resulting in stalling of replication and transcription if left unrepaired, which induces apoptosis or lethal genomic instability [73, 74]. Due to proliferation, cancer cells require much more frequent replication and transcription events, leading to greater susceptibility to platinum treatment compared to somatic cells. Many cancers, however, are either inherently resistant to the current platinum-based therapies or acquire resistance during treatment [75]. This resistance limits the range of tumors that can be treated with these platinum compounds and hinders the widespread development of fully curative treatments.

The mechanisms cancer cells use to survive treatment with platinum compounds include decreased influx, increased sequestration by intracellular thiols, and increased efflux [76]. These processes all serve to limit the amount of active platinum in the cell and thereby decrease the number of platinum-DNA lesions that form. Cancer cells can also become resistant to platinum compounds by increasing the rate at which they repair platinated DNA [77]. Inhibition of DNA platination and removal of lesions prevent the stalling of polymerases that read

DNA and the consequent induction of apoptosis [78]. Cancer cells also employ polymerases that can replicate through platinum lesions that persist or form during DNA replication, in order to prevent stalling [79]. Translesion synthesis (TLS) is a mechanism naturally employed by cells to prevent common DNA damage from stalling replication forks and giving rise to high levels of apoptosis [80, 81]. For cisplatin resistance in particular, TLS appears to be critical. Cisplatin treatment efficacy is inversely correlated to expression levels of DNA polymerase η [82]. Pol η is specialized in bypass of UV light-induced cyclobutane pyrimidine dimer lesions [32]. The enlarged active site and rigid DNA binding properties of Pol η allow it to incorporate nucleobases opposite large and helix-distorting, intrastrand crosslink adducts formed by cisplatin in a manner similar to the CPD [83]. The enlarged active site of the polymerase accommodates the crosslink and permits insertion of dC opposite the modified bases. After the first two insertion steps, however, Pol η is not proficient in extension past cisplatin lesions *in vitro*, and Pol ζ is postulated to extend the primer further until high-fidelity polymerases can rebind [32]. siRNA knockdown of Pol η or Pol ζ hypersensitized cell cultures to cisplatin, confirming that TLS plays a role in cisplatin resistance *in vitro* [84].

To overcome the resistance that gives rise to decreased efficacy of platinum therapy, compounds with alternative, non-classical molecular structures are being investigated [85]. Monofunctional platinum compounds differ from the clinically employed bifunctional species in that they form only one covalent bond to DNA [86]. Phenanthriplatin, or *cis*-diamminephenanthridinechloroplatinum(II), is a novel monofunctional platinum agent that has displayed very promising

anticancer activity (Fig. 4) [87]. Phenanthriplatin was identified as a result of a systematic variation of the *N*-heterocyclic ligand, informed by the crystal structure of RNA polymerase II stalled at a pyriplatin-platination site [88]. Phenanthriplatin maintains a spectrum of activity that is distinct from that of any other platinum agent tested in the NCI60 human tumor cell line anticancer drug screen and is 7 to 40 times more potent than cisplatin. The complex interacts covalently with DNA, presumably at the nucleophilic N7 position of guanine, and inhibits transcription by RNA polymerase II [87, 89, 90]. Phenanthriplatin contains a center of chirality and can therefore form diastereomeric adducts with DNA. Small molecule studies indicate that rotation about the bond between the platinum center and the phenanthridine ligand or the guanine ($Pt–N_P$ and $Pt–N_G$, respectively) is facile but that one diastereomeric form is preferred over the other (Fig. 4B) [91]. Monofunctional platinum adducts and phenanthriplatin in particular represent a promising new design for next generation chemotherapies but further study will be needed prior to their use.

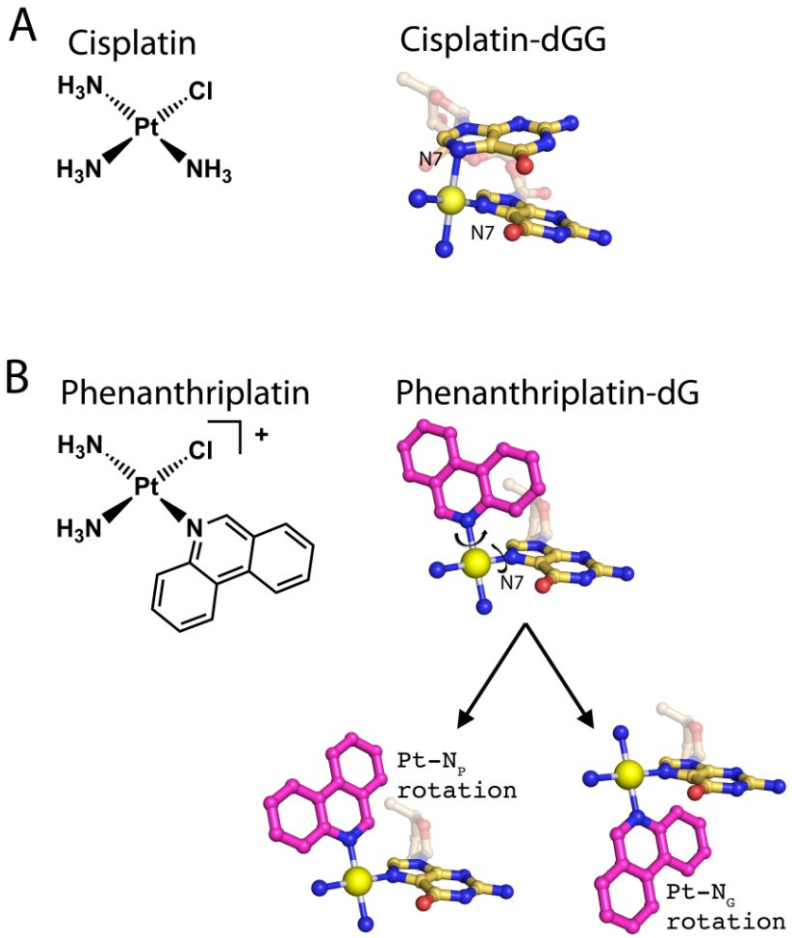


Figure 4. Chemical structures of cisplatin and phenanthriplatin and depictions of cisplatin and phenanthriplatin damaged DNA

A. The cisplatin compound binds two guanines, forming a bulky and rigid bifunctional adduct. B. Phenanthriplatin binds only a single guanine to form a more flexible monofunctional adduct. The carbon atoms of the phenanthridine ligand are shown in magenta. The major degrees of freedom available to the flexible phenanthriplatin lesion are demonstrated.

1.5. Objectives of thesis research and significance of findings

The first objective of my thesis research was to characterize the mechanism of Pol η catalysis at the molecular level, specifically concerning alignment and deprotonation of the primer. Recently a new technique, time resolved crystallography has been developed for use with Pol η[27]. This technique has enabled me to visualize atomic details of the active site in a time-lapse series during Pol η catalysis. Previous studies have implicated a number of moieties in deprotonating the primer 3´-OH, including the S113 residue and the transient water[27, 48-51]. My work identifies a key structural transition from a ground state conformation to a reactant state conformation that is critical for efficient deprotonation and catalysis, and also identifies the S113 residue as the mediator of this transition. Additionally, I collected data that demonstrated that the transient water was not solely responsible for primer deprotonation, and suggests that the general base can be varied in the polymerase catalysis mechanism. Given the highly conserved nature of polymerase active sites, these finding likely also shed light onto the mechanism used by most, if not all, members of the polymerase superfamily.

The second objective of this thesis research was characterization of Pol η as a target for inhibition. Due to Pol η's proven role as a resistance gene against chemotherapy treatment in cancers, Pol η inhibition represents a promising therapeutic goal[82]. This research was approached from two fronts: identification of small molecule binding sites for rational design of conventional inhibitory compounds, and development of platinum chemotherapeutics that present a

greater impediment to Pol η TLS. In this work I identified a novel druggable pocket in the Pol η thumb domain and verified the inhibitory potential of this binding site biochemically[32]. Further, I observed binding of dNMPNPPs at the inhibitory site in crystal structures of Pol η and confirmed the location through directed mutagenesis[32]. Though no compounds have been found yet to target this site, such drugs have to potential to increase chemotherapeutic efficacy through adjuvant therapy.

Platinum-based chemotherapies are among the most successful anti-cancer treatments available currently, but current platinum chemotherapeutics are only effective against some cancer types and are limited by acquired resistance in cancers[74, 92]. The platinum-based derivative, phenanthriplatin, shows improved cytotoxicity over traditional chemotherapeutics in a variety of cancer cell lines[87]. Here I present an investigation of the effect of phenanthriplatin adducts on replication and the ability of DNA polymerases to replicate past a site-specific phenanthriplatin lesion. Among a panel of DNA polymerases, Pol η was the only one able to bypass the phenanthriplatin lesion, although it does so with a very low efficiency. Kinetic studies of the different steps in TLS past phenanthriplatin lesions were carried out, and the results are interpreted in light of the crystal structures of the polymerase stalled at the first two steps of TLS. These structural studies reveal the nature of the interaction of Pol η with phenanthriplatin platinated DNA. They also elucidate the mechanism by which the structure of phenanthriplatin inhibits TLS in a manner distinct from, and more potent than, current platinum chemotherapeutics. This work

demonstrates the therapeutic potential of monofunctional platinum chemotherapeutics such as phenanthriplatin in overcoming known cancer resistance mechanisms and progressing toward more curative cancer treatment.

2. **Methods**

2.1. Pol η cloning and expression in E. coli

The gene encoding the human Pol η 1-432aa fragment was codon optimized for expression in *E. coli* and ordered from Genescript. The coding region was cloned into a modified pET28p plasmid using NdeI and XhoI restriction enzymes, leaving the N-terminal 6xHis tag and protease cleavage site. The pET28p plasmid is a modified version of the pET28a plasmid in which the n-terminal thrombin cleavage site is substituted for a PreScission protease site[93]. The cloned Pol η 1-432aa construct was transformed into BL21 (DE3) cells for expression. In LB broth the expression strain was grown at 37˚C to an OD~0.6-0.8 then cooled briefly on ice to ~16˚C. Expression was initiated with addition of 0.4 mM IPTG and incubated at 16˚C over night (~20 hrs) after which cells were harvested, frozen in liquid nitrogen, and stored at -80˚C until purification.

2.2. Purification of Polymerase η protein

Frozen Pol η 1-432aa expression cell pellets were thawed and suspended in cold lysis buffer (20 mM Tris pH 7.5, 1000 mM NaCl, 10 mM imidazole, 3 mM BME, Sigma protease inhibitor tablet). Cell disruption was performed with an Emulsiflex system and the lysate was clarified by ultracentrifugation at 35,000 rpm for 35 minutes followed by 0.22 µm filtration. The lysate was loaded onto a 5ml nickel column equilibrated in NickelA buffer (20 mM Tris pH 7.5, 1000 mM NaCl, 15 mM imidazole, 3 mM BME, 5% glycerol), then washed with 60 mM imidazole to remove non-specific binding proteins, and finally eluted with 300 mM imidazole. The eluted protein was buffer exchanged with a NAP-25 column into

PreScission buffer (20 mM Hepes/NaOH pH 7.0, 500 mM KCl, 3 mM DTT, 0.1 mM EDTA, 10% glycerol), cleaved at 4°C overnight with 1:100 ratio PreScission Protease, and then again buffer exchanged with a NAP-25 column into MonoS-C buffer (20 mM MES pH 6.0, 250 mM KCl, 3 mM DTT, 0.1 mM EDTA, 10% glycerol). The protein was next loaded onto a 5 ml MonoS column pre-equilibrated with MonoS-C buffer and eluted by gradient (up to 1 M KCl in 100 ml). The eluted protein was pooled, concentrated, and run on a Superdex75 column pre-equilibrated in Sup75 buffer (20 mM Tris pH 7.5, 450 mM KCl, 3 mM DTT). The purified Pol η eluted as a monomer at ~10.5ml, while a small aggregate peak is present at ~8ml elution volume. The resulting monomer peak fractions are >98% pure Pol η protein. An identical procedure was used for all Pol η mutants.

2.3. DNA Oligonucleotides

Oligonucleotides were obtained from Integrated DNA Technologies (IDT), the FDA Facility for Biotechnology Resources (FBR), or Sigma Aldrich. The oligonucleotide sequences used for steady state kinetic measurements and run-off polymerization assays are shown in their respective figures or in Table 2. Oligonucleotides used for crystallization are as follows: dA terminated primer template duplex (5´- CATTTTGACGCT -3´ and 5´- AGCGTCAA -3´), rA terminated primer template duplex (5´- CATTTTGACGCT -3´ and 5´- AGCGTCArA -3´), dT terminated primer template duplex (5´- CATTATGACGCT -3´ and 5´-AGCGTCAT -3´), 2´-fluoro-dA terminated primer template duplex (5´-

CATTTTGACGCT -3′ and 5′- AGCGTCAA$_F$ -3′), Insertion complex against phenanthriplatin (5′- CAT**G**CTCACACT -3′ and 5′- AGTGTGAG -3′), +1 extension complex against phenanthriplatin (5′- CATC**G**TCACACT -3′ and 5′- AGTGTGAC -3′). A bold **G** denotes the site of phenanthriplatin damage. Phenanthriplatin damaged DNA oligonucleotides were provided by our collaborators in the Stephen Lippard Lab at MIT.

2.4. Steady state primer extension assay

Steady state kinetic parameters K_M and K_{cat} were measured using the template/primer pairs shown in the respective figures or tables. The reaction mixture contained 2.5 or 5 nM Pol η, 0-60 μM dNTP, 5 μM 5′ 6-FAM labeled primer and template (1:1000 or greater enzyme to substrate ratio), 40 mM Tris pH 7.5, 5 mM MgCl2, 100 mM KCl, 10 mM Dithiothreitol, 0.1 mg/ml Bovine Serum Albumin, 5% glycerol. Reactions were initiated by addition of the dNTP corresponding to the next incorporation only (i.e. if template position is T then dATP is added) and 5 mM $MgCl_2$ followed by incubation at room temperature for 5 minutes. These conditions allow extension of primers only by 1 base but multiple turnovers for each enzyme. The reactions were quenched with an equal volume of formamide, 10 mM NaOH, and xylene cyanol. After heating the samples at 90°C for 3 minutes and rapid cooling on ice the samples resolved on 20% poly-acrylamide gels containing 7.5M urea at 300V. The gels were pre-run 20 minutes to increase the running temp to >30˚C, loaded with 4ul sample, and run until the Bromophenol Blue dye was nearly completely run out of the gel.

The gels were visualized using a Typhoon Trio (GE Healthcare), quantified using ImageQuantTL (GE Healthcare) software, and curve fitting was performed using Prism 5 (Graphpad) software. Representative gel and quantification is shown in figure 5 to demonstrate the separation of substrate (primer) and product (extended primer) achieved. Reported errors are calculated standard error of the mean (SEM) from triplicate experiments and best-fit error values were calculated with a least squares non-linear regression analysis in Prism 5(Graphpad). All kinetics data were fit to the Michaelis-Menten equation, $Y=Vmax*X/(Km+X)$, where Y is the extension rate, Vmax is the extension rate at infinite incoming nucleotide concentration, X is the concentration of incoming nucleotide, and Km is the Michaelis-Menten constant. The inhibition activity seen in figure 6 was fit with a one phase decay equation, $Y=(Y0 - Ymin)*exp(-K*X) + Ymin$, where Y is extension rate, Y0 is the extension rate at the highest measured value, Ymin is the asymptotic value of Y at infinite incoming nucleotide concentration, K is the rate constant, and X is the incoming nucleotide concentration.

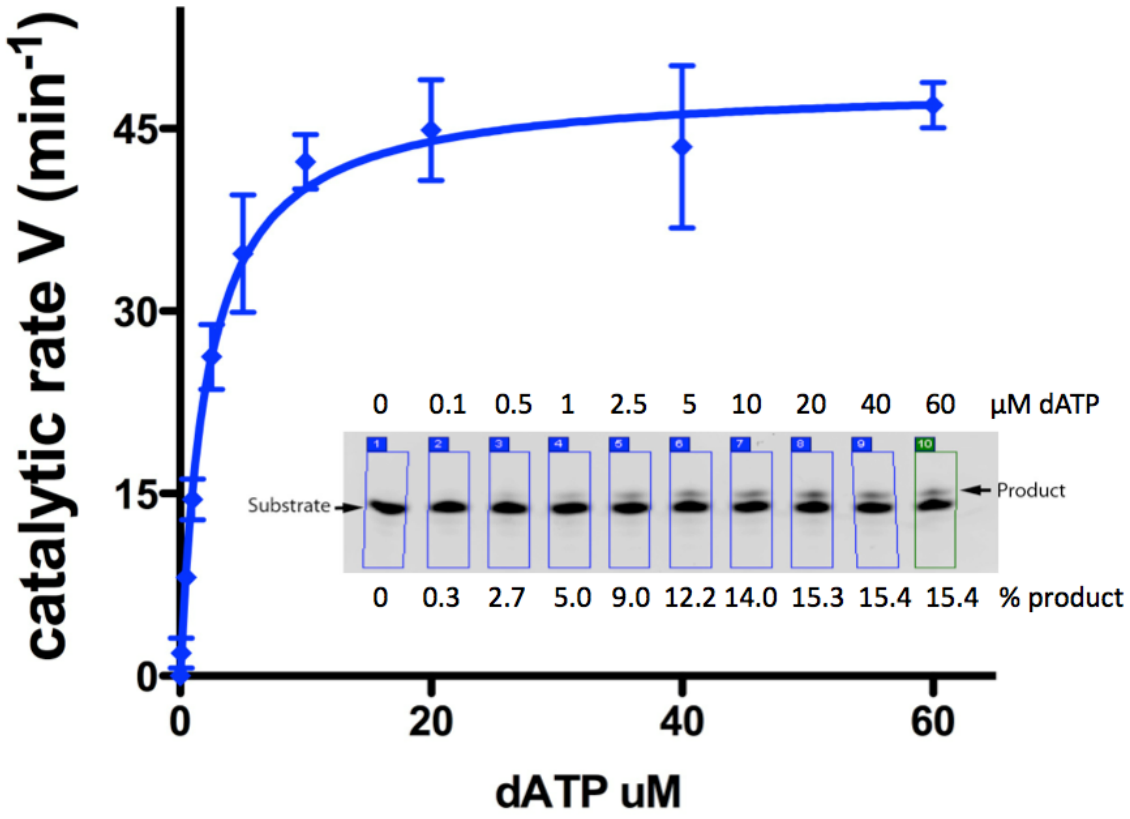


Figure 5. Representative steady state kinetics measurement

Reactions of wild type Pol η extending a dT terminated primers were run into a 20% TBE-UREA gel. Product formation was quantified with ImageQuantTL and fit to the Michaelis-Menten equation, Y=Vmax*X/(Km+X), to determine the kinetic parameters K_M and k_{cat}.

2.5. Run-off polymerization assay

Pol η 1-432aa was cloned, expressed, and purified as described in section 2.2. Pol v and Pol ζ were generously supplied by Young-Sam Lee and cloned, expressed, and purified as described previously [94]. Klenow fragment was obtained from New England Biolabs (Product# M0210S). Pol κ was obtained from Enzymax (cat# 27).

The template/primer pairs used are shown in figure 17. The reaction mixture contained 0-50 nM Polymerase, 25 μM each dNTP, 100 nM 5´ 6-FAM labeled primer and template, 40 mM Tris pH 7.5, 5 mM MgCl2, 100 mM KCl, 10 mM Dithiothreitol, 0.1 mg/ml Bovine Serum Albumin, 5% glycerol. Reactions were initiated by addition of all four dNTP (0.25 mM each) and 5 mM $MgCl_2$, incubated at 30°C for 5 minutes (10minutes for Pol ζ), and quenched with an equal volume of formamide, 10 mM NaOH, and xylene cyanol. After heating samples at 90°C for 3 minutes and rapid cooling on ice, the primers were resolved on 20% poly-acrylamide (15% for Pol ζ) gels containing 7.5M urea at 300V after a 20 minute pre-run, visualized using a Typhoon Trio (GE Healthcare), quantified using ImageQuantTL (GE Healthcare) software.

2.6. Nucleotide preference (fidelity) assay

Nucleotide preference was measured using the template/primer pairs shown in figure 18. The reaction mixture contained 2 nM Pol η, 25 μM dNTP, 5 μM 5´ 6-FAM labeled primer and template, 40 mM Tris pH 7.5, 5 mM MgCl2, 100 mM KCl, 10 mM Dithiothreitol, 0.1 mg/ml Bovine Serum Albumin, 5% glycerol. Reactions were initiated by addition 0.25 mM of the indicated dNTP and 5 mM $MgCl_2$, incubated at room temperature for 5 minutes, and quenched with an equal volume of formamide, 10 mM NaOH, and xylene cyanol. After heating at 90°C for 3 minutes and rapid cooling on ice the primers were resolved on 20% poly-acrylamide gels containing 7.5M urea at 300V after a 20 minute pre-run, visualized using a Typhoon Trio (GE Healthcare), and quantified using ImageQuantTL (GE Healthcare) software.

2.7. Crystallization and Data Collection

For crystallization, purified Pol η was mixed in a 1:1.09 molar ratio with DNA and either 5 mM $MgCl_2$ (static structures) or 1:1 $CaCl_2$ and 1:1.5 dATP (time resolved structures) followed by incubation on ice to form a binary complex. DNA oligonucleotides used are indicated for each experiment. After three-fold dilution to reduce salt concentration to 150 mM KCl the complex is concentrated to at least 3.5 mg/ml and the appropriate dNMPNPP (static structures) or dATP (time resolved structures) was added to form the ternary complex. Crystals were obtained using the hanging drop vapor diffusion method over a reservoir containing 0.1 M MES pH 6.0, 15-20% PEG 2000-MME. Microseeding was used

when consistent crystal sizes were difficult to obtain with simple vapor diffusion. For static structures of the enzyme-substrate complex, crystals were soaked briefly in 0.1 M MES pH 6.0, 20% PEG 2000-MME, and 20% glycerol for cryo protection then flash frozen in liquid nitrogen. For time resolved, reaction progress structures, crystals were soaked 30 minutes in 20% PEG 200-MME, 0.1 M Mes/KOH pH 7.0, 1 mM DTT, 5 μM dATP then transferred to 20% PEG 2K-MME, 0.1 M Mes/KOH pH 7.0, 1 mM DTT, 1 mM $MgCl_2$ for the indicated time. The crystals were then briefly transferred through 20% PEG 2K-MME, 0.1 M Mes/KOH pH 7.0, 1 mM DTT, 1 mM $MgCl_2$, 20% glycerol for cryo protection prior to freezing in liquid nitrogen. Diffraction data were collected at 100K on the 22BM beamline at the Advanced Photon Source (APS).

2.8. Macromolecular Structure Determination

Diffraction data were processed with HKL2000 or XDS and converted to structure factors by TRUNCATE [95-98]. All datasets were in the P61 space group and isomorphic with previously solved Pol η structures (PDB code 3MR2 or 4ECQ), which were used to obtain phases, using simple rigid-body refinement. Models were built in COOT and refined in PHENIX [99, 100]. All structural figures were rendered in PyMol [101]. Data collection and refinement statistics are shown in appendix 2.

3. Inhibition of DNA Polymerase η by high nucleotide concentration and a potential inhibitor binding pocket

3.1. Polymerase η has reduced activity at high incoming nucleotide concentrations

Once I had obtained sufficiently pure Pol η protein, I set out to measure the steady state kinetic constants governing its polymerization activity. A previous member of our lab had noticed inconsistent measurements for Pol η at high dNTP concentrations. I measured Pol η steady state primer extension, and visualized extended primers by gel electrophoresis. Surprisingly, when I measured Pol η's kinetic profile, focusing on high dNTP conditions, I found that Pol η's kinetic profile is different from other polymerases I have tested. The polymerization activity increases with increasing incoming nucleotide (dNTP) concentration until ~80 μM (Fig. 6). However, dNTP concentrations beyond 80 μM have an inhibitory effect on overall Pol η activity. Fitting of the low dNTP concentration phase and high dNTP concentration phase separately, with a Michaelis-Menten and one phase decay, respectively, gives an estimate of the inhibition as roughly 40 percent loss in activity at dNTP concentrations > 0.5 mM with an estimated K_d for the inhibitory binding site of 256 μM dATP.

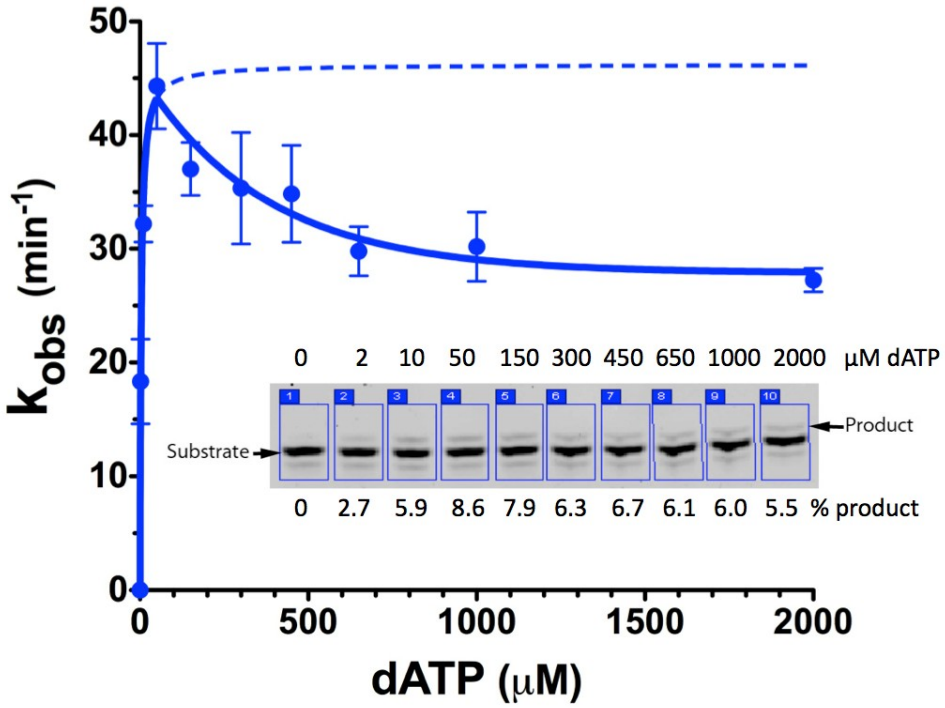


Figure 6. Inhibition of Pol η by high dATP concentration

Steady state extension activity at dATP concentrations greater than ~500 μM Pol η is inhibited (solid blue). The expected activity of Pol η in the absence of dATP dependent inhibition is shown with a dashed blue line. Data points at low dATP concentration were fit with the standard Michaelis-Menten equation, see methods, and used to extrapolate the expected activity in absence of inhibition. Data points above 50 μM dATP were fit with a one phase decay equation, see methods. Inset shows the separation of the substrate (primer) and product (extended primer) in 20% TBE-UREA gel.

The deviation of Pol η kinetics from a simple Michaelis-Menten kinetic model demonstrated that either dNTP could bind Pol η at multiple sites or bind at the active site in multiple orientations. The high conservation of active site configuration across all polymerase families suggests multiple binding motifs within a polymerase active site are not likely. The most likely explanation is that a second, lower affinity binding site for dNTP exists. Inhibition by dNTP binding at this site is not competitive, as saturation fails to eliminate activity, but is instead a case of non-competitive inhibition, explaining the reduced, but non-zero, activity at saturation.

3.2. A Secondary nucleotide binding site in the Pol η thumb domain

In searching for a secondary binding site for dNTP I re-examined previously solved Pol η ternary complex crystal structures to identify binding pockets[32]. Each structure had a small, unexplained density in an exposed hydrophobic pocket between the thumb domain and the bound DNA (Fig. 7A). All previous structures had been solved from crystals grown in buffers containing 1 mM dNMPNPP, a non-hydrolysable nucleotide analog, and the density could be explained by modeling the base moiety of the dNMPNPP used in crystallization. The ribose and phosphate moieties were disordered and not observed. The modeled dNMPNPP was stabilized by π-π interactions with tryptophan 297 residue within the thumb domain (Fig. 7A). I produced a W297A mutant to remove the proposed interaction. The mutation significantly reduced k_{cat}, likely due to destabilization of the thumb domain and DNA binding, but also

eliminated the inhibitory effect at high dNTP concentrations, returning the kinetic profile a normal Michaelis-Menten kinetics (Fig. 7B). Loss of the inhibitory activity identifies the binding site near W297 as being responsible for inhibition at high dATP concentrations. Physical identification of the binding site confirms that the inhibition is noncompetitive because the inhibitor binding site is non-overlaping and far away from the site of substrate dNTP binding.

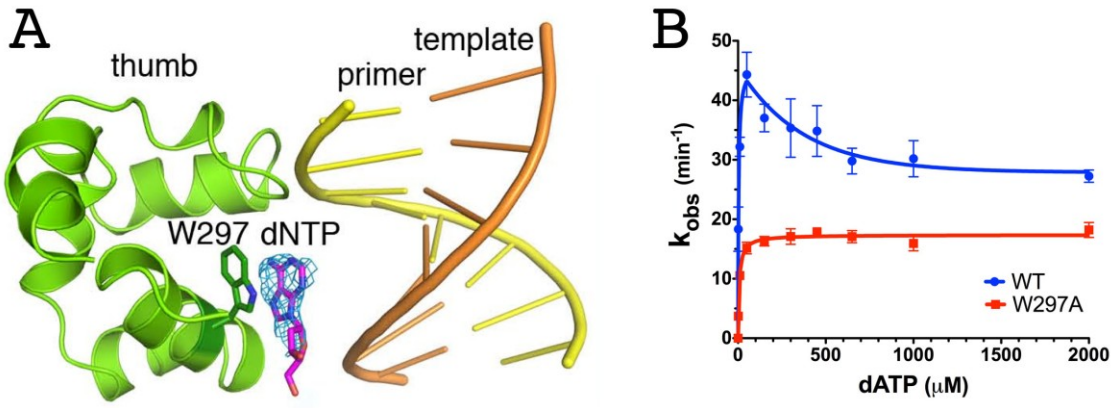


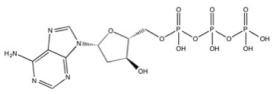
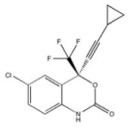
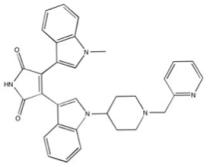
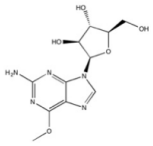
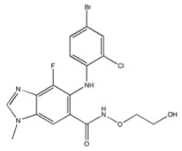
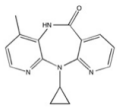
Figure 7. The inhibitory binding site is located on the thumb domain between Pol η and the upstream DNA

A. dATP bound between W297 in the thumb domain (shown in green) and the upstream DNA duplex (shown in yellow and orange). The structure is masked with an omit map of the dATP shown in blue at 2σ. B. Steady state kinetics of W297A mutant Pol η is not inhibited by high dATP concentrations, confirming dATP binding at this site is responsible for the observed inhibition of WT Pol η. These figures are adapted from my published work[32].

3.3. The Secondary nucleotide binding site is a target for potential inhibitor development

Despite Pol η's role in cancer resistance to chemotherapy no pharmaceuticals are currently available that inhibit Pol η, though there is significant interest in inhibitor development. The demonstrated inhibitory incoming nucleotide binding pocket is a promising target for rational drug design. The structure of an incoming nucleotide offers a starting point for design of inhibitor candidates. As only the base portion of the nucleotide is sufficiently ordered to be visible in the structures, compounds with large aromatic rings offer the most promise. I have screened a number of aromatic compounds available to my lab but, aside from dATP, none have been effective (Table 1). A wider library of aromatic compounds will be needed to identify specific higher affinity inhibitors. Among polymerases, compounds capable of binding this site will likely be quite specific for Pol η, as sequence comparison with other Y-family DNA polymerases reveals that the W297 site is unique to Pol η (Fig. 8).

Table 1. Screen for aromatic Pol η inhibitors

Compound	Structure	Inihibition	Compound	Structure	Inihibition
dATP		~40%	efavirenz		None
Enzastaurin		None	Nelarabine		None
Selumetinib		None	Nevirapine		None

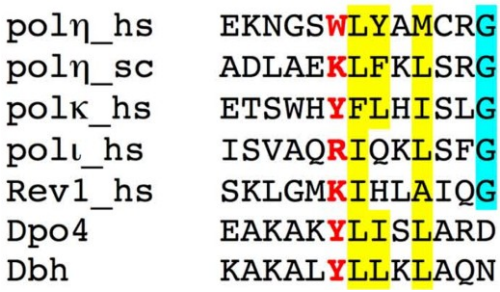


Figure 8. Sequence alignment of the Pol η inhibitory dATP binding site

Y-family polymerases in human (hs) and yeast (ys) aligned around the helix containing W297A. W297A and the corresponding positions from other polymerases are shown in red. Conserved hydrophobic residues that form the pocket are shown highlighted in yellow, non-hydrophobic conserved residues in cyan. This figure is reprinted from my published work[32].

4. The active site S113 residue facilitates alignment of the DNA primer end for catalysis

4.1. Sequence conservation within Polymerase η active site

The active site of Y-family polymerases (residues contacting the primer terminus, magnesium ions, and incoming nucleotide) resides at the interface of the palm and finger domains. Both domains have a number of highly conserved residues whose functions have been elucidated with structures of a variety of polymerases[102]. The palm domain contains the three critical acidic residues that coordinate two magnesium ions required for catalysis[25]. The magnesium ions are also coordinated by the triphosphate of the incoming dNTP and balance its large negative charge. The finger domain also interacts with the incoming nucleotide with conserved polar and aromatic residues that stabilize its binding in the correct orientation for catalysis[102].

Of the conserved residues near the active site nearly all have been assigned roles as either important or required for stabilizing the protein fold, structural integrity, DNA binding, or polymerase reaction activity[102]. However, the polymerase tertiary structures used in assigning these roles were obtained using modified substrates or polymerase mutants to inhibit activity to gain homogeneity in sample preparation and facilitate crystallization. The most common modification is to remove the 3´-OH from the terminal base of the primer[23, 103, 104]. This hydroxyl, once deprotonated, acts as the nucleophile during catalysis and its removal completely inactivates the polymerase/DNA complex. The lack of this important component has hindered assignment of catalytic roles to some active site residues, such as the S113 position in human Pol η.

Within the Pol η active site S113 is positioned very near the catalytic 3´-OH of the primer strand and E116, an acidic residue that coordinates the A-site magnesium ion (Fig. 9B). S113 is highly conserved within the Y-family polymerase family[102] (Fig. 9A). The analogous positions in the A- and B-family polymerases are strictly conserved, though the amino acid conserved within each polymerase family is different, histidine and threonine, respectively (Fig. 9B-D). Conservation of this position suggests an important catalytic role but thus far none has been demonstrated.

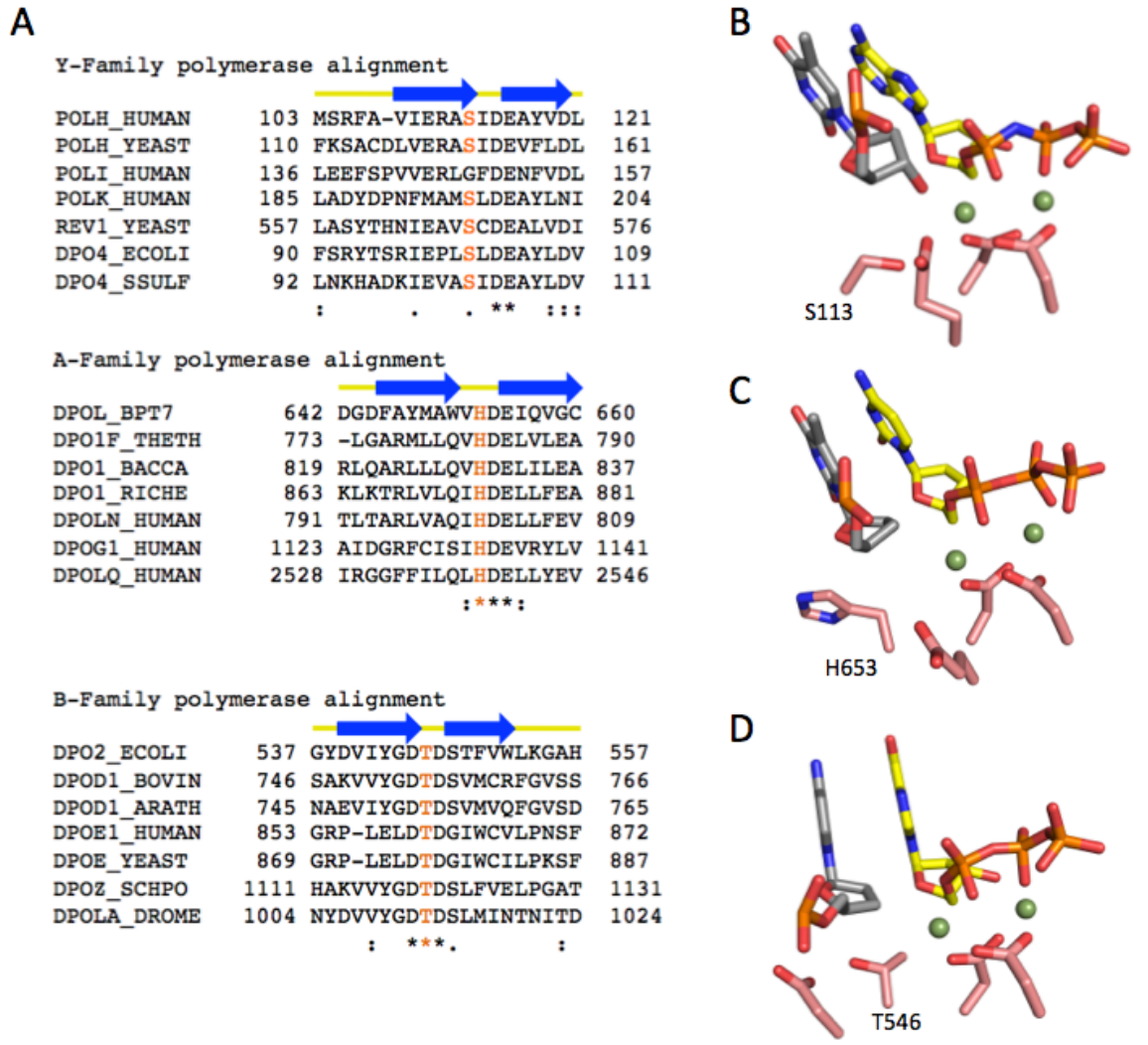


Figure 9. Alignment of active site residues within the polymerase sequence motif III

A. A-family alignment includes: Bacteriophage T7 Pol (DPOL_BPT7, P00581), T. *thermophiles* Pol I (DPO1F_THETH, P30313), B. *caldotenax* Pol I (DPO1_BACCA, Q04957), R. *Helvetica* Pol I (DPO1_RICHE, Q9RLB6), H. *sapiens* Pol v (DPOLN_HUMAN, Q7Z5Q5), H. *sapiens* Pol G1 (DPOG1_HUMAN, P54098), H. *sapiens* Pol Q (DPOLQ_HUMAN, O75417). B-family alignment includes: E. *coli* Pol II (DPO2_ECOLI, P21189), B. *taurus* Pol δ (DPOD1_BOVIN,

P28339), A. *thaliana* Pol δ (DPOD1_ARATH, Q9LVN7), H. *sapiens* Pol ε (DPOE1_HUMAN, Q07864), S. *cerevisiae* Pol ε (DPOE_YEAST, P21951), S. *pombe* Pol ζ (DPOZ_SCHPO, Q9P6L6), D. *melanogaster* Pol α (DPOLA_DROME, P26019). Y-family alignment includes: H. *sapiens* Pol η (POLH_HUMAN, Q9Y253), S. *cerevisiae* Pol η (POLH_YEAST, Q04049), H. *sapiens* Pol ι (POLI_HUMAN, Q9UNA4), H. *sapiens* Pol κ (POLK_HUMAN, Q9UBT6), S. *cerevisiae* Rev1 (REV1_YEAST, P12689), E. *coli* Dpo4 (DPO4_ECOLI, Q47155), S. *solfataricus* Dpo4 (DPO4_SULSF, P96022). The structurally conserved S113 position in human Pol η is highlighted in each alignment. The conserved secondary structures are also shown above the sequences. Alignment was performed with the Clustal Ω program[105]. B. Active site of Y-family DNA Polymerase η demonstrating the structural position of the S113 residue (pdb code, 3MR2)[25]. C. Active site of A-family T7 DNA polymerase demonstrating the similarity of the H653 position to that of the S113 residue of the Y-family DNA polymerase η, however H653 is pointing in away from the active site, likely due to the lack of a primer 3´-OH used in crystallization(pdb code, 1T8E)[103]. D. Active site of B-family DNA polymerase II demonstrating the structural similarity of the T546 position to that of the S113 residue of the Y-family DNA polymerase η (pdb code, 3MAQ)[104].

4.2. Mutation of Polymerase η S113 reduces incorporation rate

Due to S113's proximity to the primer 3´-OH, I hypothesized that S113 is a player in the deprotonation process, which is still not understood for the polymerase reaction. An S113A mutant has made previously[27] and I used this mutant to assess roles that S113 may play in catalysis. The serine to alanine mutation was chosen to remove the hydrogen bond and acid/base chemistry competent hydroxyl, while maintaining the general size and structure of the residue[27]. The steady state kinetic parameters of single nucleotide incorporation were measured for the mutant extending a dA terminated primer . The S113A mutant has an 8- fold higher K_M and 5-fold lower catalytic activity than WT (Fig. 10A and Table 2). The S113A mutant activity has been measured previously on dA terminated primers by Nakamura et al[27]. Their measurement of the S113A mutant's activity was considerably higher, having 95% of the k_{cat} value seen for the WT. However, I believe their published measurements are inaccurate as I have not been able to reproduce this high activity. Using new protein and reagents or using samples of protein and reagents obtained from the original authors, the S113A mutant activity is consistently ~5 fold lower than WT, having a k_{cat} of only 13 min^{-1}. If S113 were the general base responsible for primer deprotonation the catalysis would be eliminated. Despite the lower activity level than previously reported, there remains observable catalytic activity for the S113A mutant which demonstrates that S113 is not essential for extracting the proton from the primer 3´-OH, consistent with the conclusions drawn by Nakamura et al[27]. However, loss of the S113 hydroxyl in the mutant does lead to

a substantial loss of activity indicating an important role, either in stabilizing the fold or in a specific hydrogen bond interaction with substrates or other active site components.

To investigate which roles S113 may be playing, the structure of the S113A mutant Pol η bound to a DNA template primer duplex, Mg^{2+} and a non-hydrolysable dAMPNPP was solved to 2.27Å (Fig. 10B). The mutant showed no gross defect in folding with an RMSD of only 0.218 Å over 399 pairs of Cα atoms when superimposed with the WT structure [25]. The alanine adopts a phi/psi angles and sidechain positioning nearly identical to that of the WT serine. Locally, the active site fold of the mutant is also unchanged, coordinating both Mg^{2+} ions and positioning the primer 3´-OH optimally (Fig. 10B). The static molecular structure alone, showing no significant perturbation, was not able to explain the large loss in activity observed in solution.

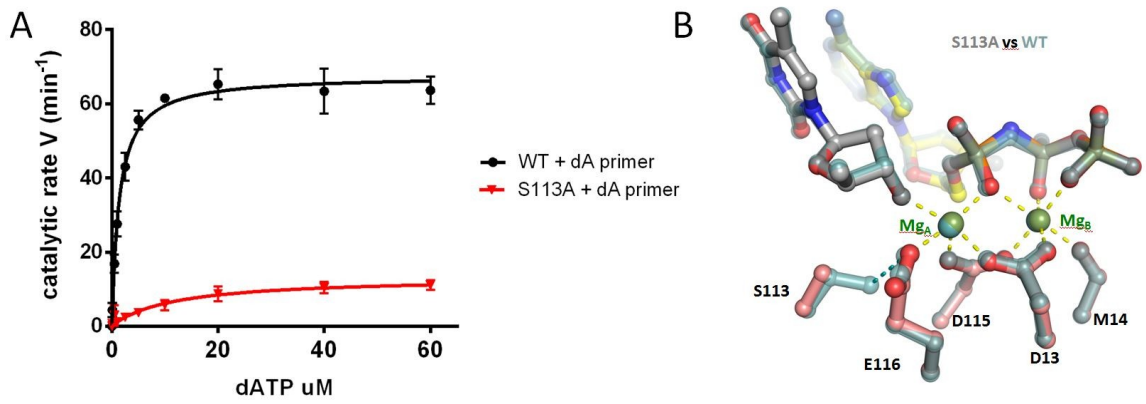


Figure 10. Pol η S113A mutant has reduced activity

A. Steady state kinetic profiles of WT and S113A mutant Pol η extending primers terminating in dA. Kinetic parameters for each profile are listed in table 2. B. Structural alignment of WT and S113A active sites. Both WT (pdb code 3MR2[25], shown in lightblue sticks) and S113A mutant (shown in colored sticks, yellow incoming dAMPNPP, salmon protein residues, and grey primer) are bound to a dT terminated primer. Mg^{2+} coordination is indicated with yellow dashes.

Table 2. Steady state kinetic parameters of WT and S113A Pol η extension of DNA and RNA terminated primers

DNA sequence	Experiment		k_{cat} [min^{-1}]	K_M [μM]	Catalytic efficiency [μM^{-1} min^{-1}]	Efficiency relative to Deoxy
3′ CACGGATCGCATTTGTACTGAG 5′ TGCCTAGCGTA**A** or 3′ CACGGATCGCATTTGTACTGAG 5′ GTGCCTAGCGTA**A**	WT	Deoxy	67.7±1.1	1.4±0.1	47.6	
		Ribose	72.8±2.3	2.0±0.3	36.4	0.76
		2'Fluoro	95.1±1.0	5.2±0.2	18.3	0.38
	S113A	Deoxy	13.3±1.3	11.5±3.3	1.2	
		Ribose	79.5±3.1	8.4±1.2	9.5	8.18
		2'Fluoro	109.2±1.8	21.4±0.9	5.1	4.41
3′ CACGGATCGCATATGTACTGAG 5′ TGCCTAGCGTA**T**	WT	Deoxy	46.8±0.9	2.1±0.2	22.3	
		Ribose	38.1±1.0	4.3±0.4	8.9	0.40
	S113A	Deoxy	ND	ND	ND	
		Ribose	43.0±1.6	7.7±0.9	5.6	ND

4.3. Non-productive conformations of the primer 3´ terminal dT observed in the S113A mutant Polymerase η active site

Static crystal structures only observe conformations that are present at equilibrium within the crystallization drop. To observe changes in transient conformations and structural transitions undergone during Pol η catalysis due to S113A mutation, I employed time resolved crystallography with S113A mutant Pol η and primers terminated in either dT or dA. Briefly, crystals of the S113A mutant Pol η were grown in complex with a primer/template duplex in the presence of equimolar calcium and excess dATP. Once fully grown the crystals were activated by transition to a magnesium containing buffer and frozen in liquid nitrogen at a number of reaction time points.

In contrast to the static structure, the S113A active site bound DNA with a dT terminated primer is not configured in a single, ground state (GS) like state in the time resolved structures at the 0 sec time point, prior to activation by Mg^{2+} addition (Fig. 11A). At 0 sec of the WT time course, only a single Ca^{2+} ion is bound in the B-site and the terminal dT of the primer strand is found uniformly in the GS position. In contrast, the ground state of the S113A mutant displayed at least two conformations[27], one conformation corresponds to the standard GS, while in the other conformation the terminal dT of the primer is shifted out of the active site and adopting a non-productive conformation (Fig. 11A). The non-productive conformation of the primer terminal base is no longer base stacked with the neighboring base in the primer stand, though it is still able to hydrogen bond with the templating base. The incoming nucleotide binding site is in steric conflict with the non-productive primer conformation resulting in reduced occupancy of the dAMPNPP and active site Ca^{2+} ion in the pre-reaction, 0 sec structure. Alternate conformations are also observed for the conserved aspartate residues responsible for Mg^{2+} binding.

The non-productive conformation of the primer terminus persists after activation of the polymerase by Mg^{2+} addition. Similar to WT Pol η, Mg^{2+} binding is observed at the metal ion A-site immediately after soaking the crystals in a buffer containing Mg^{2+}, though the A-site occupancy is only 30%. Despite observed Mg^{2+} binding in structures solved at time points up to 480 sec no bond formation is detectable, from either of the two primer conformations (Fig. 11A). The subpopulation of complexes with non-productive primer conformations

cannot react due to the lack of bound dATP. However, dATP is bound within the subpopulation of Pol η complexes (50%) with the primer terminus in the normal GS position, but even this subpopulation shows no activity suggesting the S113A mutation is defective for more than just orienting the primer prior to Mg^{2+} binding.

The steric clash between the incoming nucleotide and the non-productive primer conformation predicts that the conformation of the active site can be modulated by concentration of incoming nucleotide. The static structure of Pol η bound to the dT terminated primer was solved in 5 mM Mg^{2+} and 100 μM dAMPNPP, which more closely mimics the concentrations *in vivo* and the solution kinetic measurements. As shown in figure 10B, at 1 mM dAMPNPP the S113A active site bound the incoming dAMPNPP and Mg^{2+} ions indistinguishably from WT. However, at the 100 μM concentration the occupancy of dAMPNPP in the drops significantly to 0.3 indicating that the mutant active site binds incoming nucleotide much weaker than the WT. The major conformation the primer terminal dT is in the non-productive conformation, the conserved active site acidic residues are shifted out of the active site, and there is reduced occupancy for the Mg^{2+} bound at both sites (Fig. 11B). The WT Pol η complex structure solved in identical, *in vivo* mimic conditions retained full dAMPNPP binding and a well organized, catalytically competent active site, demonstrating that the S113A mutation reduces the affinity for the incoming nucleotide when the primer is not stabilized in the GS position by a hydrogen bond with the S113 hydroxyl. The loss of the S113 to 3´-OH interaction likely contributes to the ~8 fold increase in K_M values seen for the S113A mutant compared to WT (Fig. 10A).

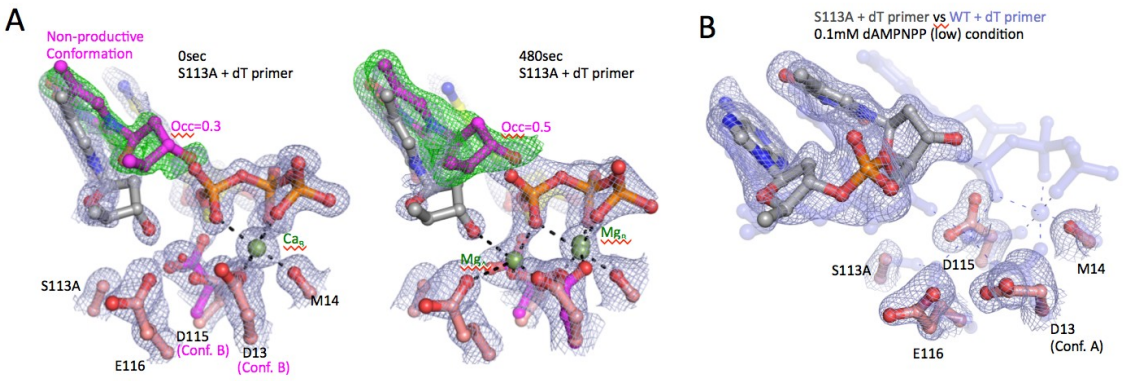


Figure 11. Non-productive conformation of the primer terminal dT base in the S113A Pol η active site

A. 2Fo-Fc (1σ) and Fo-Fc (3σ) maps calculated omitting the non-productive primer conformation superimposed onto the 0 sec (left) and 480 sec (right) structures of S113A Pol η - dT terminated primer complex. The non-productive primer and active site residue conformations are shown in magenta. B. Static structures of WT (shown in semi-transparent blue) and the major conformation of S113A (shown in grey and salmon) Pol η under physiological concentrations of incoming nucleotide, 100 μM dAMPNPP. 2Fo-Fc density of the S113A refinement masks the structure.

4.4. Reaction competent alignment of the primer is influenced by base stacking of the primer 3´ terminal base

To investigate roles of S113 in DNA synthesis by Pol η, time resolved structures of S113A mutant Pol η extending a primer terminated with dA were solved. The adenine base, which is known for the best stacking potential[104], at the primer terminus almost completely eliminates the non-productive primer conformation, and the canonical GS conformation is observed at 0 sec (Fig. 12). This is likely due to the increased base stacking between the primer strand and the incoming nucleotide that stabilizes the 3' end of the primer strand[106]. Specifically, the primer terminal adenine base stacks with the upstream primer base prior to dNTP or Mg^{2+} binding, stabilizing the productive conformation sufficiently to eliminate non-productive conformation observed with pyrimidine terminated primers (Fig. 11), even without the S113 to 3´-OH interaction. Productive binding of the adenine terminated primer in turn allows for efficient dNTP and Mg^{2+} binding, as seen in time courses with S113A mutant on dA terminate primers (Fig. 12). Base stacking of primer terminal pyrimidine bases is not sufficiently strong to stabilize the productive primer conformation in the absence of the S113 to 3´-OH interaction.

Proper binding and alignment of the primer terminus is thus influenced by both the base stacking potential and hydrogen bonding of the 3´-OH with the active site S113 residue. S113 stabilization appears to be more important for extension of pyrimidines than purines, due to their weak base stacking potential, and allows for more sequence-independent extension rate. The more ordered

dA terminated primer can be extended by the S113A polymerase *in crystallo*. Upon activation with Mg^{2+} exposure, the metal A-site is occupied and the reaction proceeds (Fig. 12). This is in agreement with solution kinetics showing the S113A mutant remains catalytically active on dA terminated primers. On dT terminated primers there is no detectable reaction occurring within the crystal, perhaps due to the decreased stability of the terminal pyrimidine nucleotide and the sub-optimal *in crystallo* reaction conditions. Steady state kinetic measurements have not yet been accurately collected in solution for S113A mutant Pol η extending a dT terminated primer, but initial experiments indicate the activity in solution is significantly lower than WT. The reduced catalytic rate of S113A Pol η on dA terminated primers, despite proper dATP binding and primer conformation, suggests S113 is involved in an additional event prior to catalysis.

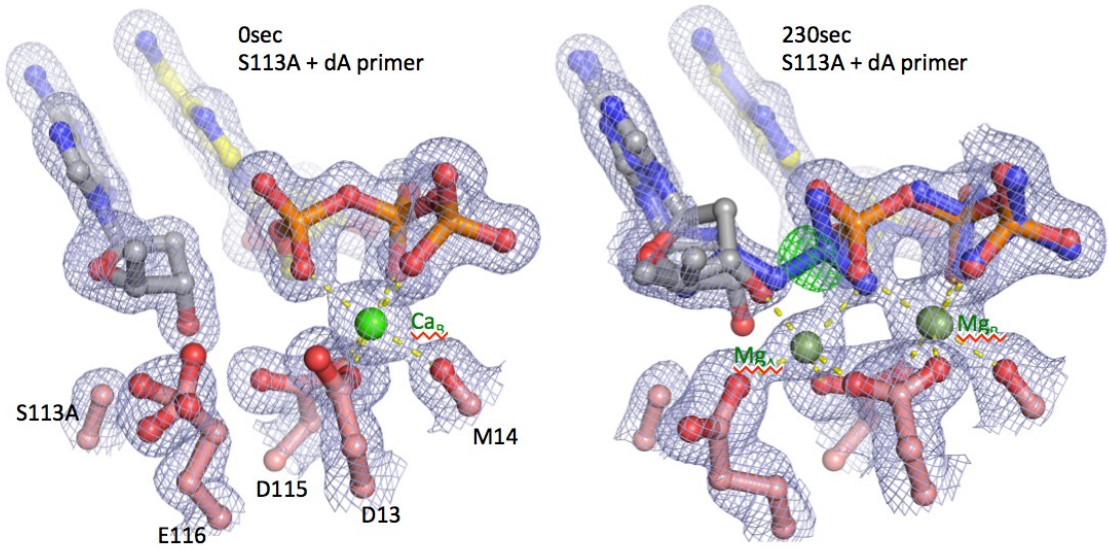


Figure 12. Active site conformation and reactivity of a dA terminated primer within S113A Pol η

Left: 2Fo-Fc (1σ) superimposed on the 0 sec S113A Pol η - dA terminated primer structure. Right: 2Fo-Fc (1σ) and Fo-Fc (3σ) maps calculated omitting the product conformation (shown in blue) superimposed on the 230s S113A Pol η - dA terminated primer structure. Density for the new bond is present between the primer and dATP.

4.5. The ground state to reactant state transition is slower in S113A mutants than in wild type Polymerase η

Loss of activity for the S113A mutant in solution cannot be fully explained by unstable binding of the primer, since dA terminated primers, which are stable in the S113A Pol η active site, still result in activity loss in solution. In good agreement with the solution kinetics, the S113A mutant reacted slower *in crystallo* than WT. Three major changes occur prior to bond formation that can account for reduced rate: Mg^{2+} binding, GS to RS transition, and primer deprotonation. Omit maps for the O3´ and C3´ atoms of the terminal base, calculated for the early time points of a WT and S113A Pol η extending a dA terminated primer, showed a marked delay in movement of the 3´-OH from the GS position to the reaction competent RS position (Fig. 13A and B). WT Pol η has a significant RS population at the 40 sec time point, while the S113A mutant has no RS population until 140sec. Additionally, the GS to RS transition is never fully complete in the S113A mutant, showing a significant GS population persisting beyond 5 minutes. The delayed GS to RS transition in the S113A mutant results in reduced overall catalytic rate, as measured in solution. The transition step from a GS to RS complex was only recently observed in Pol η structures [27]. In the previously reported WT Pol η *in crystallo* reaction, the RS was formed rapidly and a delay was observed between the RS appearance and the beginning of product formation. Thus, for WT Pol η the GS to RS transition is not considered to be the rate-limiting step[27]. However, for the S113A mutant, the GS to RS transition appears to be the slowest step in our *in crystallo* reaction,

suggesting that one role S113 serves is to accelerate the transition of primer end to the reaction-ready state. This role is likely accomplished by the switch of the S113 hydrogen bond interaction from the 3´-OH to E116 upon binding of Mg^{2+} to the A-site seen in WT time courses [27].

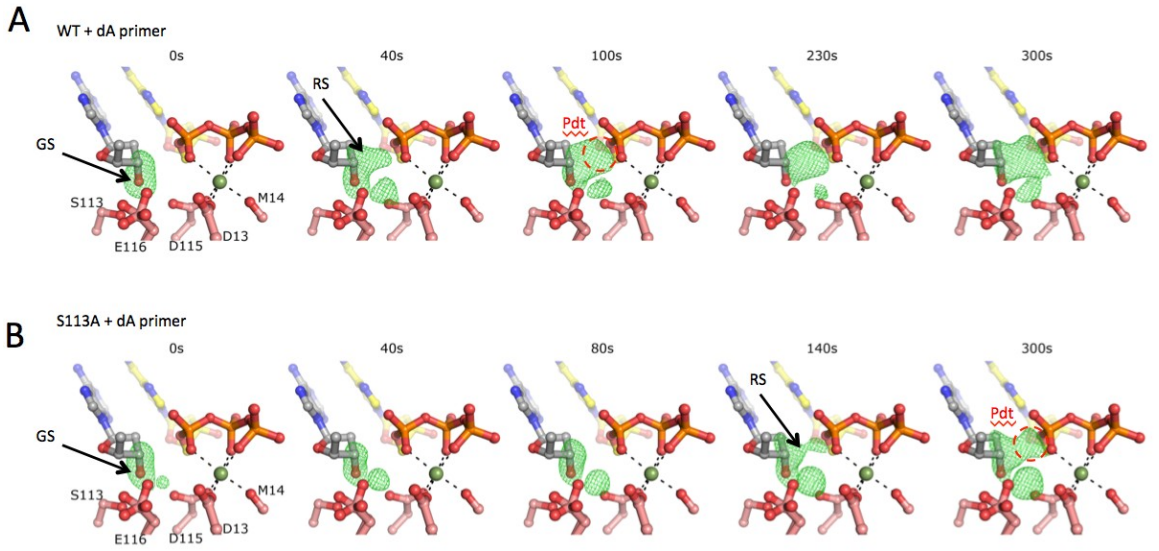


Figure 13. The ground state to reactant state transition rate is reduced by S113A mutation

A. $Fo_{(WT\ 0sec-300sec)}$-$Fc_{(WT\ 0sec)}$ omit maps (3.5σ) of the C3´ and O3´ atoms superimposed onto the 0 sec WT Pol - dA terminated primer complex. GS and RS positions are indicated with arrows. Product density is circled in red.

B. $Fo_{(S113A\ 0sec-300sec)}$-$Fc_{(WT\ 0sec)}$ omit maps (3.5σ) of the C3´ and O3´ atoms superimposed onto the 0 sec WT Pol - dA terminated primer complex. GS and RS positions are indicated with arrows. Product density is circled in red.

4.6. RNA terminated primers rescue S113A Polymerase η activity to wild type levels by eliminating the requirement for a ground state to reactant state transition

To investigate the effect of the GS to RS transition rate on the overall activity, primers terminated with RNA bases were used to measure steady state extension rates in solution. Deoxyribonucleotides in a double helix have a strong bias for C2´-endo pucker of the ribose group, while ribonucleotides in a double helix exist exclusively in the C3´-endo pucker conformation [107]. Extension of RNA terminated primers is catalyzed by WT Pol η at a rate comparable to DNA primers. For S113A mutant Pol η, RNA terminated primers restore its activity to near WT levels (Fig. 14A). Both normal RNA bases and modified RNA bases containing fluorine at the 2´ position were able to rescue the defect, indicating that the identity of the 2´ group was not important for the rescue. The static structure of S113A Pol η bound to an rA terminated primer with Mg^{2+} and dAMPNPP was solved. Consistent with the known sugar pucker bias, the terminal primer RNA base adopts a C3´-endo pucker and the base undergoes a slight propeller twist but the active site remains otherwise unaffected (Fig. 14B). This indicates that the altered sugar pucker is responsible for the rescue of S113A activity by removing a slow kinetic step prior to formation of the RS, seen at equilibrium in the static structure.

Time resolved *in crystallo* structures were collected of S113A mutant Pol η extending an rA terminated primer, to visualize the effect of the sugar pucker change on the GS to RS transition. Unlike the dA terminated primer, prior to Mg^{2+}

addition, the 3´-OH of the rA terminated primer is positioned in the RS conformation, despite no Mg^{2+} being bound in the A-site (Fig. 14C). Throughout the reaction, no GS population was detected (Fig. 14D). Due to the sugar pucker bias of the RNA ribose ring, the 3´-OH is occluded from the GS conformation, forcing it to adopt the RS position. Thus, the GS to RS transition is no longer required, alleviating the kinetically slow GS-to-RS transition in the mutant and resulting in restored activity. The RS and C3´-endo positioned primer did appear to reduce the incoming nucleotide affinity, resulting in reduced occupancy for the dATP, but did not affect the reaction in fully bound complexes. The S113A activity on RNA primers is only rescued back to the WT level, suggesting the GS to RS transition is no longer rate-limiting.

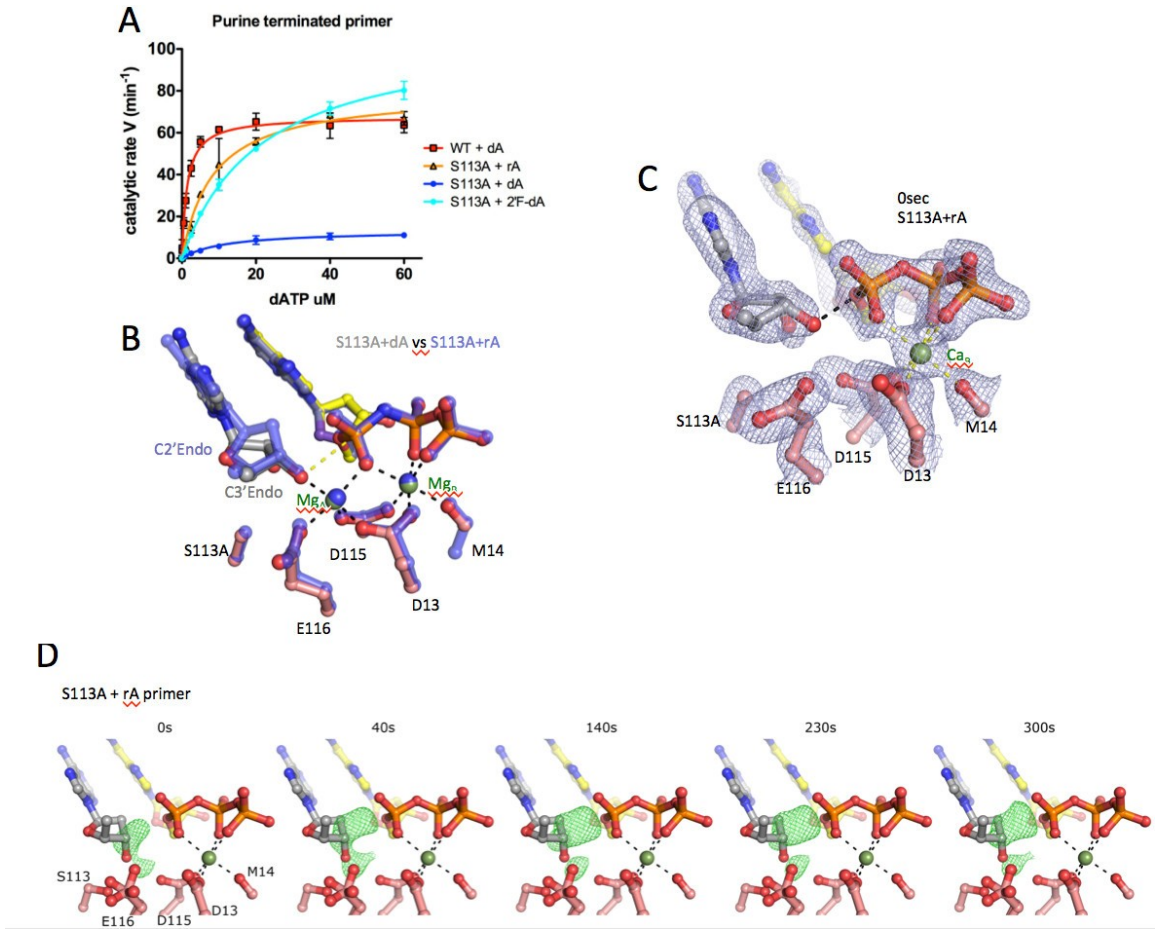


Figure 14. The intrinsic C3'-endo sugar pucker bias of RNA terminated primers rescues S113A Pol η activity

A. Steady state kinetic profiles of WT and S113A mutant Pol η extending primers terminating in dA/rA/2´fluoro-dA. Kinetic parameters for each profile are listed in table 2. B. Structural alignment of S113A Pol η active sites bound to dA or rA terminated primers. The dA primer structure is shown in semi-transparent blue sticks and the rA primer structure is shown in colored sticks (yellow incoming dAMPNPP, salmon protein residues, and grey primer). The pucker change of

the terminal primer base is labeled. C. 0 sec structure of S113A Pol η bound to an rA terminated primer and masked with 2Fo-Fc map (1σ). The rA primer is bound in the RS position prior to A-site Mg^{2+} addition/binding. D. Fo(S113A + rA primer 0sec-300sec)-Fc(WT + dA primer 0sec) omit maps (3.5σ) of the C3´ and O3´ atoms superimposed onto the 0 sec WT Pol η - dA terminated primer complex. GS and RS positions are indicated with arrows. Product density is circled in red.

4.7. Deprotonation is not dependent on the transiently 3´-OH associated water observed in time-resolved Pol η structures

The time resolved structures of WT Pol η extension reported by Nakamura et al showed that two Mg^{2+} ions are essential to align the primer 3´-OH with regard to the α-phosphate of the incoming dNTP in the RS [27]. Formation of the RS took place within 40 sec after Mg^{2+} addition. However, there is a significant time delay after both Mg^{2+} ions are bound (RS), during which the reactants seem to be perfectly aligned but before the nucleophilic attack and new phosphodiester bond take place, beginning at 80sec. The delay suggests that a slow chemical change, which may involve subtle conformational changes, is required to reach the transition state and complete catalysis. This slow step appears to be the rate-limiting step of WT Pol η catalysis. Deprotonation of the 3´-OH is a prerequisite for the nucleophilic attack, and could be the rate-limiting step that corresponds to the delay between binding of Mg^{2+} ions and phosphodiester bond formation[45]. The RS delay ends when density appears for a water molecule that

hydrogen bonds directly to the 3´-OH group, termed the transient water because it disappears due to a steric clash after new phosphodiester bond formation[27] (Fig. 15A). The transient water appearance correlates temporally with phosphodiester bond formation, suggesting that it may be the general base that extracts the proton from the 3´-OH or serve to accept the 3´-OH proton once otherwise prompted to disassociate. Steitz has previously proposed that the two Mg^{2+} ions play the key role of the general base[28]. Since the 3´-OH is directly coordinating the A-site Mg^{2+} ion in the active site, its pKa can be dramatically reduced, promoting deprotonation. However, Mg^{2+} can only prompt the disassociation of the proton; it cannot accept the displaced proton. The transient water is positioned well to play the role of proton acceptor (Fig. 15A).

To investigate the role of the transient water in deprotonation of the primer, the structure of WT Pol η bound to a rA terminated primer and dNMPNPP was solved. The 2´-OH moiety occupies a position in the RS conformation that occludes the transient water site (Fig. 15A). Despite the absence of the transient water, the reaction proceeds in solution at a nearly unchanged rate, indicating the deprotonation mechanism was not significantly disrupted by the loss of the transient water. However, the 2´-OH, occupying the transient water site, is also competent to participate in proton shuttle chemistry. Using a water-mediated proton transfer mechanism, in which protons rearrange their bonding within a hydrogen bond network, the 2´-OH could accept the 3´-OH proton while transferring its proton to solvent. Consistent with such a transfer, a water is seen in structures with RNA terminated primers that hydrogen bonds to the 2´-

OH (Fig. 15A). The modified RNA base 2´-fluoro-dA was tested as a more effective block of proton transfer through the transient water site, because the fluorine atom is not a proton acceptor. In solution Pol η extends 2´-fluoro-dA terminated primers efficiently, demonstrating the transient water, though temporally correlated to bond formation, is not necessary for catalysis or primer deprotonation /shuttling (Fig. 15B).

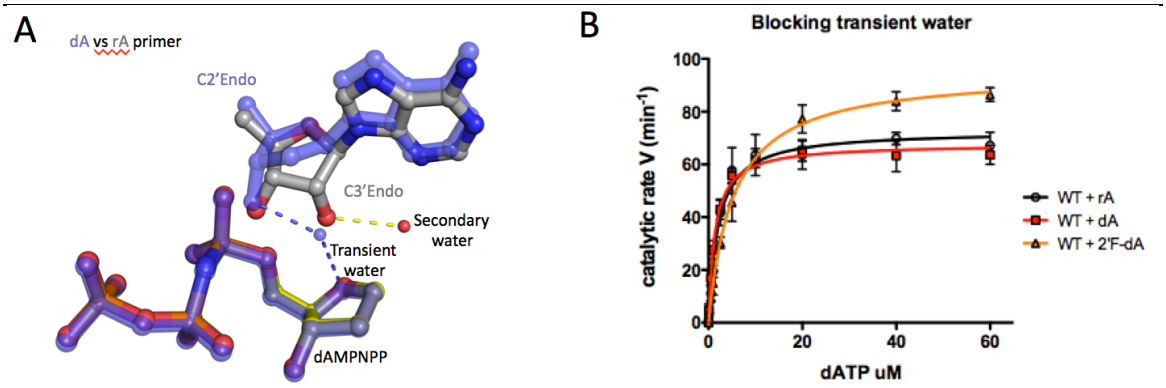


Figure 15. Primer deprotonation during Pol η catalysis does not require the transient water

A. Structural alignment of the static Reactant State structures of WT Pol η bound to dA (semi-transparent blue sticks) and rA (grey and yellow sticks) terminated primers. The transient water location is shown as a semi-transparent sphere. The secondary water of the rA structure is shown as a red sphere. Hydrogen bonds between the terminal primer sugar groups and the waters are shown with dashes. B. Steady state kinetic profiles of WT Pol η extending primers terminated in dA/rA/2´fluoro-dA.

4.8. The proton departure path may be variable

I have shown that neither the S113 residue nor the transient water is necessary to accept the primer proton during deprotonation (Fig. 10A and 13B). To identify other possible proton carriers the static structure of WT Pol η bound to the 2´-fluoro-dA terminated primer with Mg^{2+} and dAMPNPP was solved to 1.31Å. The high resolution ensures that even weakly bound groups and solvent will be observable. Candidate proton carriers were identified using the criteria that they must be as least proximal enough to the 3´-OH to engage in a hydrogen bond, 3.2Å or less. Four groups were 3.2Å or closer to the primer 3´-OH: D115 (Oδ), E116 (Oε), dAMPNPP (αO), and a water that directly coordinates the A-site Mg^{2+} (Fig. 16). The non-bridging oxygen of the α-phosphate is coordinating the A-site Mg^{2+} ion and has an extremely low intrinsic pKa value, making it a very unlikely proton acceptor [108]. The acidic residues D115 and E116 are also directly coordinating the A-site Mg^{2+} and thus have lowered pKa values due to coordination of the Mg^{2+} ion[53], similar to the primer. The residue homologous to E116 in DNA Pol β has been implicated as the proton acceptor during polymerase catalysis, and consistent with this conclusion its mutation reduces activity dramatically [33, 48]. However, due to E116's role in coordination of the catalytic Mg^{2+}, mutational approaches are unable to unambiguously demonstrate that activity loss is due to a deprotonation defect rather than disruption of proper Mg^{2+} coordination. The final candidate, the water proximal to the 3´-OH, is the only solvent ligand of the A-site Mg^{2+} ion. This A-site Mg^{2+} water has been previously proposed to directly accept the 3´-OH proton in the mechanism of

Dpo4 and Pol β using QM/MM simulations [45, 50]. Though this water is directly coordinating Mg^{2+} as well, it is capable of rearranging its O-H bonds with other waters with which it hydrogen bonds, allowing for continued coordination of the catalytic Mg^{2+} while shuttling the proton out to bulk solvent via a water mediated proton transfer mechanism. Blocking of this site is not experimentally tractable, so this proton path also cannot be unambiguously verified.

Efforts to identify the proton acceptor in polymerase catalysis have not converged onto a consensus location. I propose that this is due to the nature of the polymerase two metal ion catalysis mechanism. The orientation of the primer proximal to the Mg^{2+} cluster favors the deprotonated 3´O state, providing the energy required to deprotonate the primer[35, 53, 109]. The proton acceptor is thus not required to be a strong general base, such as those used in other acid/base catalysis mechanisms. Instead, the polymerase deprotonation pathway may not use one fixed proton acceptor but instead be capable of using multiple groups, either residues or solvent, as has been proposed in QM/MM simulations of Pol β [45]. Unfortunately, redundant routes for the proton out of the active site impede identification by traditional mutagenesis. A variable proton acceptor and path would account for the inability to identify a singular general base for polymerase catalysis thus far, despite considerable effort in the field.

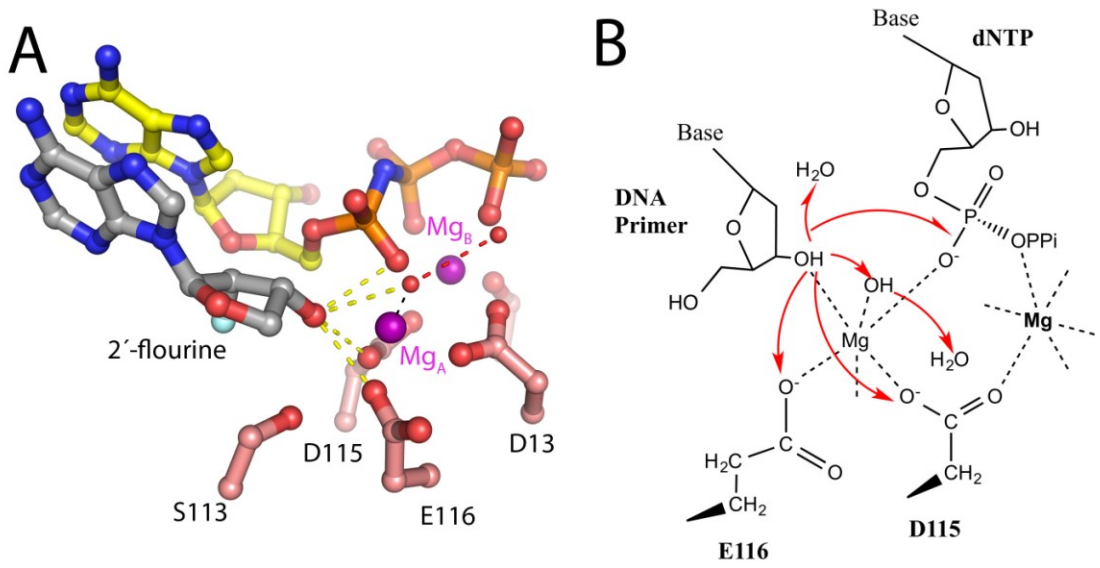


Figure 16. Identification of alternate deprotonation routes for the 3´OH proton

A. The structure of WT Pol η bound to 2´-fluoro-dA terminated primer is shown with the primer in grey, dAMPNPP in yellow, and protein residues in salmon. Yellow dashes connect the 3´-OH to all proximal proton acceptors (<3.2Å). The red dash indicates the hydrogen bond network available at the A-site coordinating water for water mediated proton transport. B. Diagram of possible deprotonation routes in the case of a 2´-fluoro terminated primer.

5. Phenanthriplatin, a novel chemotherapeutic candidate, induces cytotoxicity through potent inhibition of both replicative and TLS DNA polymerases

5.1. Phenanthriplatin-dG lesions stall replicative and TLS DNA polymerases at the extension step of lesion bypass

To investigate the effects of phenanthriplatin on DNA synthesis, I first determined the phenanthriplatin bypass efficiency of a replicative DNA polymerase (Klenow fragment) and a variety of human TLS polymerases. The TLS polymerases studied included the A-family Pol ν, B-family Pol ζ, and Y-family Pol κ and Pol η (Fig. 17). Collaborators in the Lippard lab at MIT generously supplied template DNA where only a single dG that was adducted with phenanthriplatin to form a single adduct for bypass. Each polymerase was able to catalyze the insertion step and incorporate a nucleotide opposite the phenanthriplatin-dG adduct with an apparent rate comparable to that obtained using undamaged DNA. Pol ν, ζ, κ, and the Klenow fragment all failed, however, to incorporate a nucleotide after the phenanthriplatin lesion and largely stalled at the +1 extension step. Only Pol η was able to catalyze the +1 extension with sufficient efficiency to fully bypass phenanthriplatin DNA lesions.

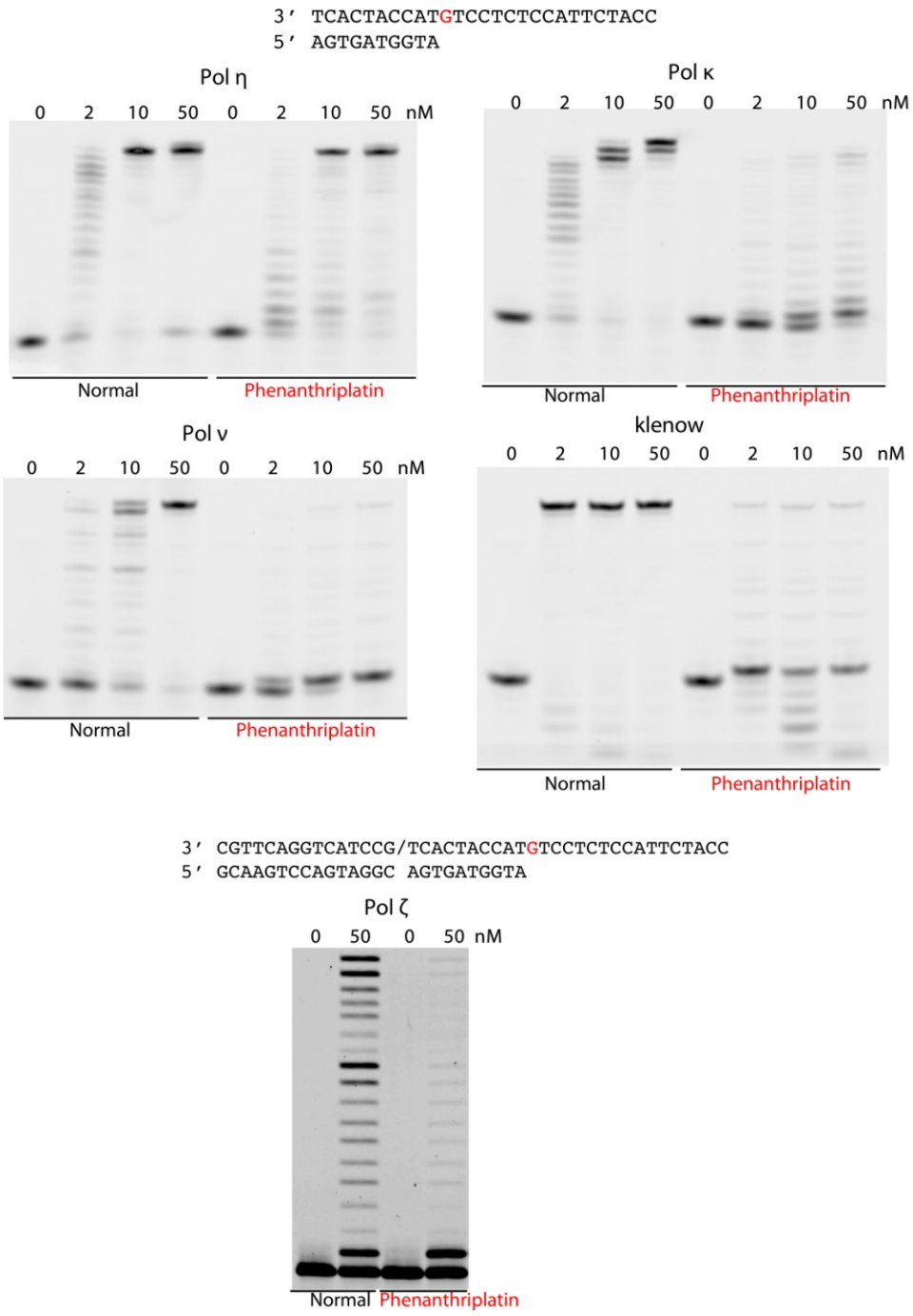


Figure 17. Comparison of run-off primer extension assays of Pol η, κ, ν, or the Klenow fragment using undamaged or phenanthriplatin damaged DNA

The DNA substrates are shown, the damage site is colored red, and the "/" indicates a nick. Polymerase concentrations used in the gradients were 0, 2, 10, and 50 nM. For Pol ζ the concentrations used were 0 and 50nM.

5.2. Fidelity of incorporation opposite phenanthriplatin-dG

In addition to the efficiency of TLS, the fidelity of bypass is also important to maintain genomic stability. TLS polymerases are often specialized to bypass their corresponding lesions with the highest fidelity possible[80]. For example, during bypass of two crosslinked thymine bases (CPD) on the template strand, Pol η preferentially incorporates two dATPs opposite the adduct. Some mutations do still occur, however Pol η reduces the mutation rate due to CPD damage [110]. For phenanthriplatin, all polymerases tested are capable of inserting opposite the damaged dG. The fidelity of insertion was measured for each of the polymerases using a nucleotide preference assay (Fig. 18A). The fidelity of insertion was very high for all polymerases tested. In fact, Pol η appears to have the lowest fidelity at the insertion step, likely due to the openness of the Pol η active site. This result suggests that the adducted dG is still able to bind in a nearly normal position within the active sites of these polymerases and form a Watson-Crick base pair with the incoming nucleotide. Given the size of the phenanthriplatin adduct, correct positioning of the damaged dG in the templating base position may requires considerable rearrangement of the phenanthridine ring moiety (Fig. 4B). Consistent with this theory, nucleotide incorporation by

high fidelity replicative polymerases opposite the bifunctional and sterically constrained cisplatin adduct is stalled [25].

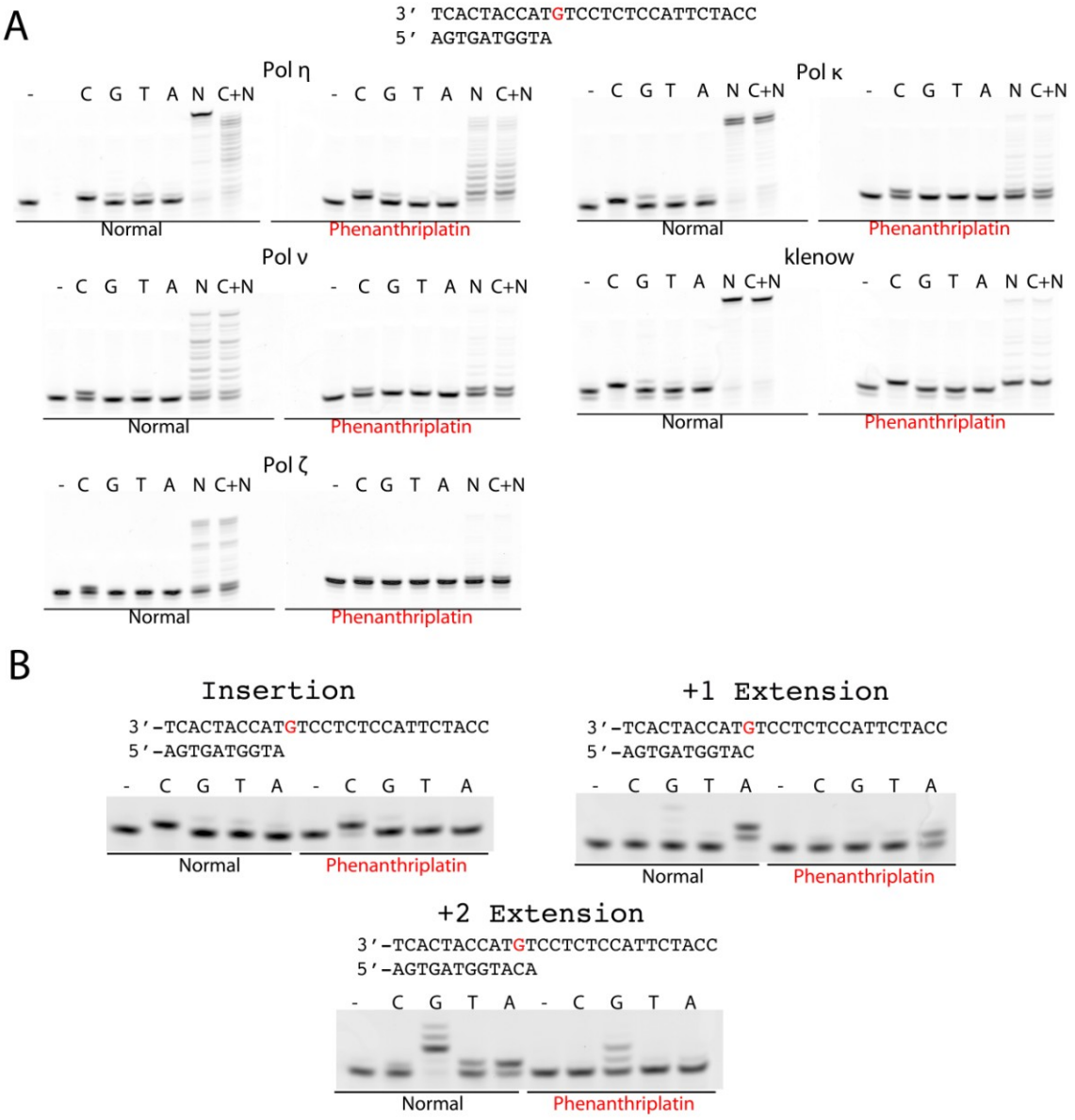


Figure 18. Fidelity of TLS bypass of a phenanthriplatin-dG adduct

A. Fidelity of Pol η, Pol ζ, Pol ν, Pol κ, and the Klenow fragment bypassing of phenanthriplatin-dG at the nucleotide insertion step. B. Fidelity of Pol η bypass of phenanthriplatin-dG in the insertion, +1 extension, and +2 extension steps.

5.3. Polymerase η bypasses phenanthriplatin-dG at low efficiency

The catalytic efficiency and fidelity of Pol η during the first three incorporation steps of TLS were determined by using a DNA oligonucleotide with a phenanthriplatin-dG adduct and normal DNA of identical sequence (sequences shown in Table 3). Under steady state conditions primers were extended one nucleotide by Pol η, and the extended primers were visualized by gel electrophoresis and fit to the Michaelis-Menten equation to determine steady state kinetic parameters. Pol η accurately incorporated dC opposite the phenanthriplatin-dG with a respectable catalytic efficiency (k_{cat}/K_M), which is 35% of that obtained with undamaged DNA (Fig. 18B, Table 3). The catalytic efficiency dropped to 5% and 4% at the +1 and +2 extension steps, respectively, for DNA platinated with phenanthriplatin. The large reduction in efficiency of the +1 extension step using platinated DNA is due to a ~6-fold increase in the K_M, indicating disruption of dNTP binding. The +2 extension step with platinated DNA displayed a 6-fold reduction of k_{cat} and a 4-fold increase of K_M, as compared to those values obtained when observing normal DNA extension. As a result, this step has a relative efficiency similar to that of the +1 extension step. Despite the overall reduced efficiency of Pol η in primer extension after the phenanthriplatin-dG lesion, its fidelity during both extension steps remained high (Fig. 18B). The absence of stalled intermediates beyond the +2 extension step in the run-off assays using platinated DNA (Fig. 17) suggests that the inhibitory effects of

phenanthriplatin begin to diminish after the +2 extension step because the lesion is translocated farther upstream from the active site.

The kinetic profile of phenanthriplatin TLS by Pol η is reminiscent of cisplatin TLS [32], in that Pol η becomes rapidly less efficient in the extension steps (Table 3). At each step, the efficiency of phenanthriplatin bypass by Pol η is about half that of cisplatin bypass. In contrast to the efficient extension beyond cisplatin adducts by Pol ζ, here Pol ζ was stalled by phenanthriplatin after insertion of a single nucleotide and failed to efficiently extend the primer (Fig. 17). It has been proposed for cisplatin bypass that Pol η is replaced by Pol ζ during the extension steps, which improves the efficiency of these steps [32, 84, 111]. Improved efficiency of the extension steps, due to exchange of Pol η for Pol ζ, leaves the second insertion effectively the lowest efficiency step of cisplatin bypass, having 47% of the efficiency observed for incorporation using undamaged DNA [32]. Pol η appears to be the only DNA polymerase capable of full TLS past the phenanthriplatin-dG adduct. No higher efficiency polymerase can replace Pol η in the manner proposed to occur in the bypass of cisplatin. Thus, phenanthriplatin bypass has two consecutive low catalytic efficiency steps, each with 4-5% of the normal efficiency, which may combine to increase the toxicity of this compound over that of cisplatin [87].

Table 3. Steady state kinetic measurements for each step of Pol η bypass of phenanthriplatin adducted DNA

Polymerization step	substrate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($\mu M^{-1}min^{-1}$)	Relative efficiency	Relative efficiency on Cisplatin
Insertion	Normal	169.4±4.6	6.8±0.6	24.9		
	Phenanthriplatin	38.5±1.6	4.5±0.7	8.6	0.35	0.59 and 0.47
+1 Extension	Normal	116.4±2.7	4.1±2.7	28.4		
	Phenanthriplatin	35.0±2.6	25.4±4.2	1.4	0.05	0.12
+2 Extension	Normal	62.9±1.6	0.8±0.1	78.6		
	Phenanthriplatin	10.5±0.3	3.1±0.3	3.4	0.04	0.08

Note: Cisplatin data reported from Zhao et al. 2012. PNAS[32].

Insertion DNA	+1 Extension DNA	+2 Extension DNA
3´-TCACTACCATGTCCTGTCCATTCTACC	3´-TCACTACCATGTCCTGTCCATTCTACC	3´-TCACTACCATGTCCTGTCCATTCTACC
5´-AGTGATGGTA	5´-AGTGATGGTAC	5´-AGTGATGGTACA

6. Structural features unique to Pol η enable TLS of phenanthriplatin-dG

6.1. Polymerase η structure in the insertion step is not affected by phenanthriplatin damage

The crystal structure of the ternary complex of Pol η (1-432aa), DNA with phenanthriplatin-dG at the templating position, and a non-hydrolysable dCMPNPP nucleotide was determined at 1.55 Å resolution. The insertion complex structure is virtually superimposable on that of an undamaged structure (pdb code 4DL3) with a root-mean-squared deviation (RMSD) of only 0.33 Å over 399 pairs of Cα atoms (Fig. 19) [32].

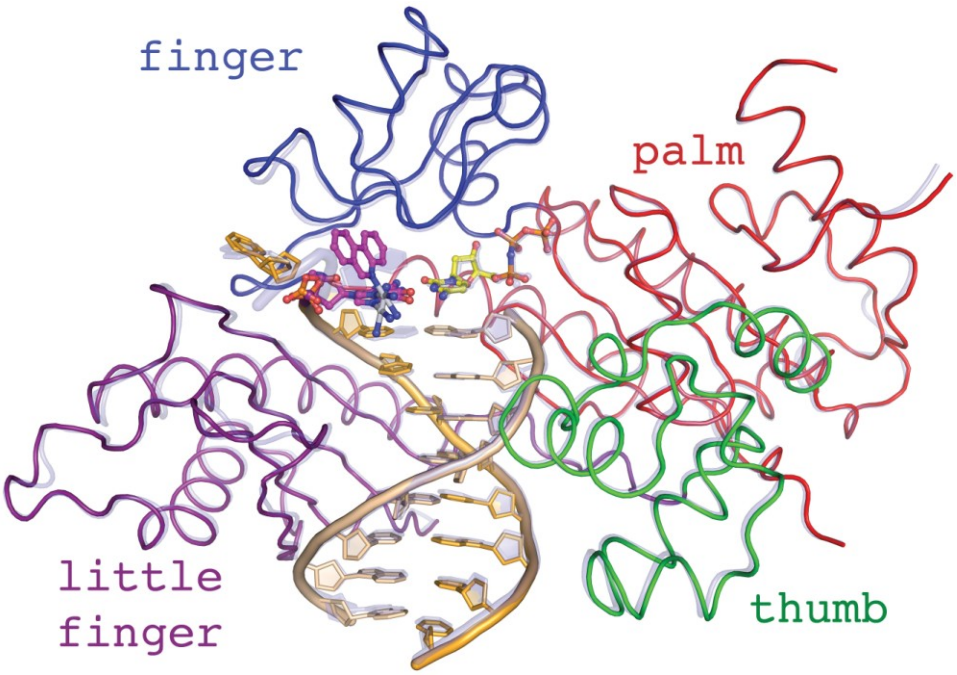


Figure 19. Superposition of the Pol η phenanthriplatin insertion structure upon the structure of Pol η bound to undamaged DNA (4DL3)

The phenanthriplatin-dG is shown in magenta. Protein and DNA from the undamaged structure are show as semi-transparent blue for comparison with the insertion complex.

6.2. The CPD-binding pocket of Polymerase η accommodates the phenanthriplatin adduct during the insertion step

In the insertion complex, the phenanthriplatin modified dG base forms a canonical Watson-Crick base pair with the incoming dCMPNPP (Fig. 20A). Two conformations of the phenanthriplatin-dG lesion are discernible. The electron density for the Pt atom is observed for both conformations, but the aromatic rings of phenanthriplatin are only observed for the major conformation with weak density (Fig. 20B). The major conformation of the phenanthridine ligand corresponds to the same conformational isomer that was observed to form preferentially in small molecule models of the phenanthriplatin-DNA lesion [91]. The two conformations of the templating base are related by a ~15° propeller rotation about the base-pair plane, with ~70% occupancy for the major and ~30% for the minor species. Phenanthriplatin forms a covalent bond to the N7 of dG in the major groove. The phenanthridine ligand is oriented toward the 5´ end of the template strand and interacts with the finger domain of Pol η. Pol η has a pocket surrounding the templating base that accommodates UV-induced cyclobutane pyrimidine dimers (CPD's) or the downstream base of an undamaged DNA template strand [25]. This pocket is used in the cisplatin bypass mechanism to accommodate the 5´-dG of the cisplatin cross-linked guanine nucleotides [32]. The present structure reveals that, during phenanthriplatin TLS, the pocket can accommodate the phenanthriplatin adduct during formation of the insertion complex (Fig. 20B).

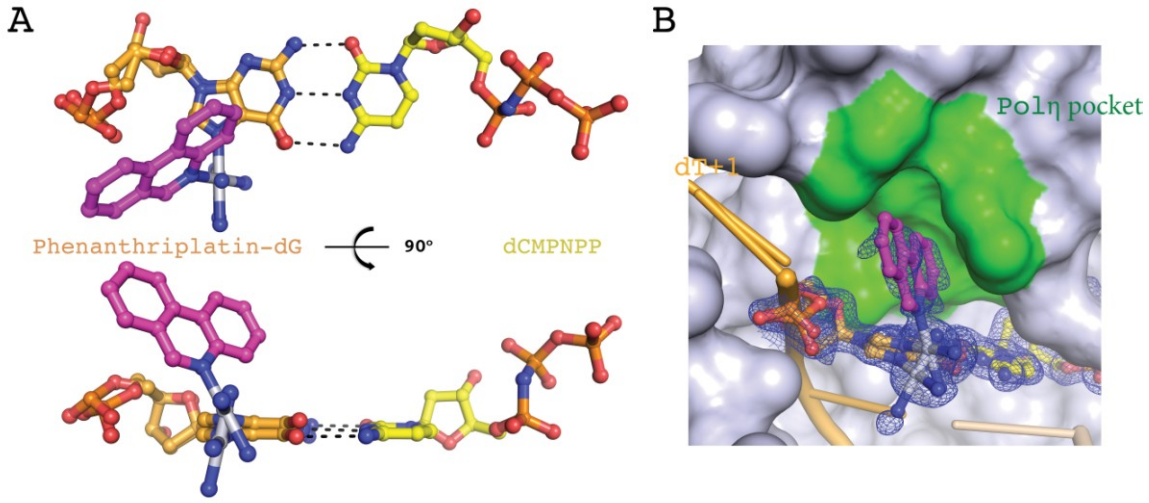


Figure 20. Pol η accommodation of phenanthriplatin adduct in the insertion step complex

A. The templating base pair. The incoming nucleotide is shown in yellow, the phenanthriplatin damaged dG is shown in orange, and the phenanthriplatin is shown in magenta. Watson-Crick base pairing interactions are illustrated with dashed lines. B. Phenanthriplatin binding pocket. The phenanthridine ligand fits into the CPD-binding pocket in the finger domain shown in green. The blue $2F_o$-F_c at 1.0 σ masks the damaged dG and incoming nucleotide.

6.3. Flexibility and conformational bias of phenanthriplatin-dG

Monofunctional platinum adducts that form a covalent bond to a single nucleotide, such as the one formed by phenanthriplatin, are substantially more flexible than CPDs or cisplatin adducts, which crosslink two adjacent bases. Moreover, rotation about Pt–N_P and Pt–N_G is facile [91]. This high flexibility explains the weak electron density observed for the phenanthridine moiety, because the binding pocket in Pol η is large enough to permit ~10° of rotation about Pt–N_G and ~20° of rotation about Pt–N_P (Fig. 21). The flexibility of phenanthriplatin may also allow it to reposition so as to be accommodated by other polymerases during the insertion step. This hypothesis would explain the activity observed for Pol ν, ζ, κ, and the Klenow fragment in the insertion step, which would require nearly a 180° rotation about the Pt–N_G bond (Fig. 18A).

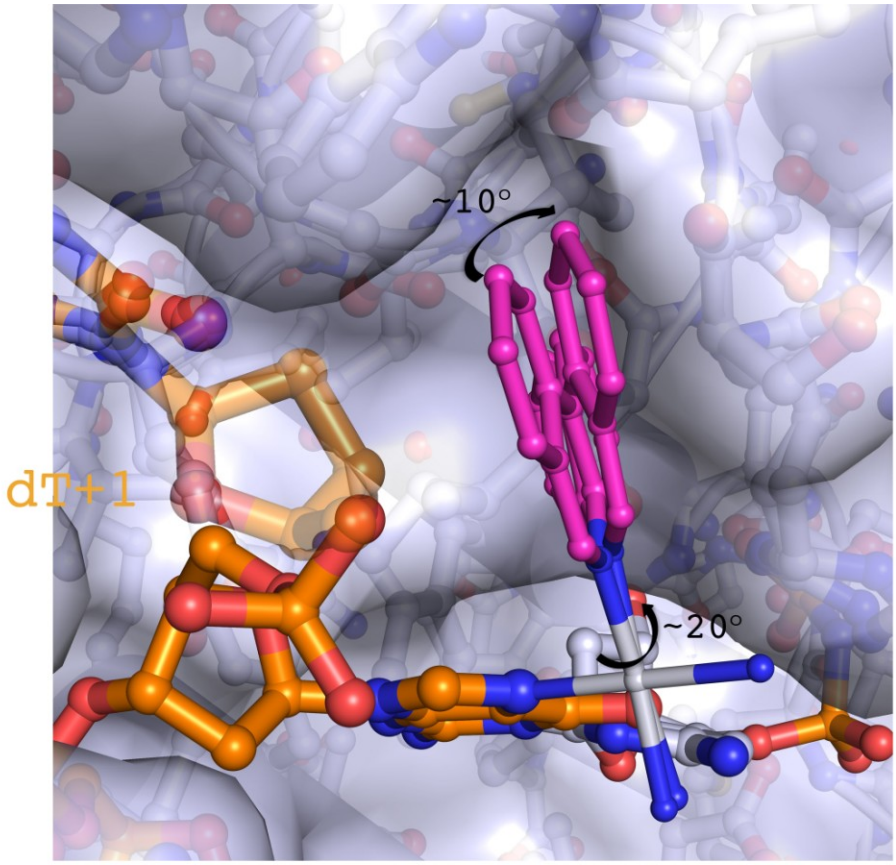


Figure 21. Model of the flexible range of phenanthriplatin within the CPD-binding pocket of Pol η

The rotational extremes possible for the phenanthridine ligand are shown in magenta. The polyhydrocarbon rings are able to rotate ~10° about the Pt–N_G bond and ~20° about the Pt-N_P bond.

6.4. Polymerase η-phenanthriplatin damaged DNA complex in the extension step reveals large conformational changes in the downstream DNA

The crystal structure of the ternary complex of Pol η (1-432aa), DNA platinated at a dG base-paired with the 3´ end of the primer strand, and a non-hydrolysable dGMPNPP base-paired with the dC downstream (+1) of the lesion, was determined at 2.8 Å resolution. Consistent with studies of Pol η TLS of other bulky DNA lesions, the protein structure in the phenanthriplatin +1 extension step remains relatively unchanged by the adduct, resulting in a pairwise Cα RMSD of 0.28 Å from an undamaged structure[25, 32]. A 2.9° rotation of the Little Finger domain away from the catalytic core is observed together with an adjustment of the template strand (Fig. 23A). The DNA, however, undergoes large conformational changes near the lesion to accommodate phenanthriplatin during the +1 extension step (Fig. 22 and 26).

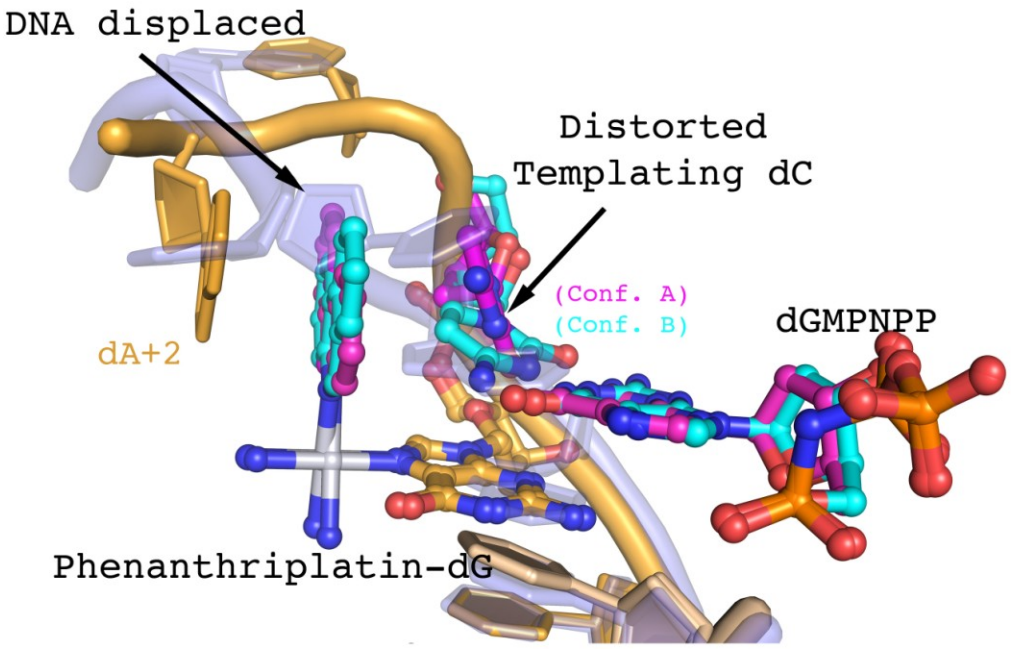


Figure 22. Structure of Pol η - phenanthriplatin-dG DNA complex at the +1 extension step

Undamaged DNA from a normal Pol η ternary complex structure (4DL3) is superimposed onto the +1 extension complex and shown in semi-transparent blue to demonstrate the DNA conformational changes. Displacement of the downstream DNA backbone and distortion of the templating dC are indicated with arrows. Two conformations of phenanthriplatin, the templating base, and the incoming nucleotide are present. The major conformation has 80% occupancy and is shown in magenta. The minor conformation has 20% occupancy and is shown in cyan.

The phenanthriplatin-dG and downstream DNA of the template strand are no longer in the canonical B-form, but the upstream duplex DNA retains B-form character because of extensive interactions with the little finger domain of Pol η. This phenomenon is known as the molecular splint effect. As evident in CPD TLS structures, a critical β-sheet involving amino acids 316-324 runs parallel to the template strand and forms hydrogen bonds between the three phosphate groups immediately upstream of the platinum adduct and every other main-chain amide (Fig. 23B) [25]. In the +1 extension complex, Arg, Lys, and Thr side chains also hydrogen bond with the phosphosugar backbone of the template strand and reinforces B-form structure. The distortion induced by the phenanthriplatin lesion is mainly absorbed by rotation of the ribose rings of the upstream bases and a 2.9° rotation of the little finger domain, leaving base pairing undisturbed.

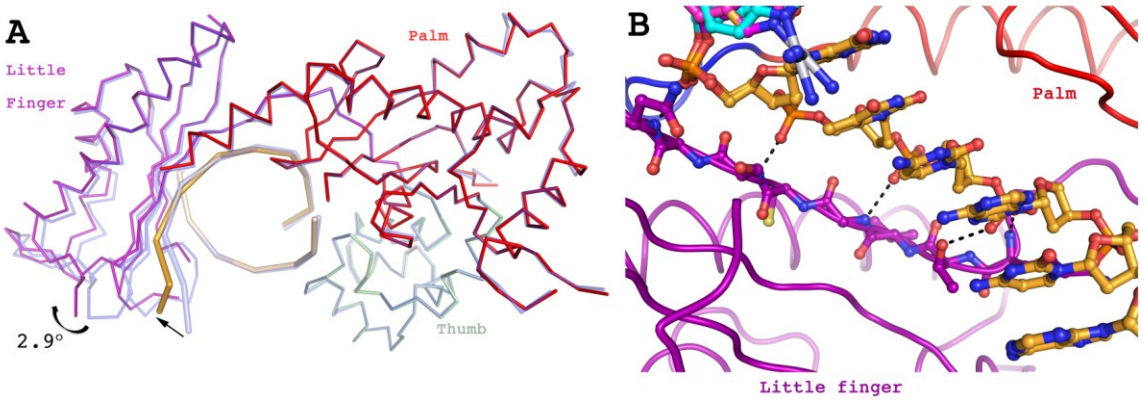


Figure 23. The molecular splint of Pol η stabilizes the damaged upstream template DNA in B-form helical structure

A. The +1 extension complex structure superimposed with the normal Pol η complex (pdb code 4DL3). The extension complex is colored by domain and the undamaged structure is semitransparent blue. The shift in DNA position downstream of the adduct is shown with a straight arrow and the rotation of the Little Finger domain is denoted with curved arrow. The finger domain is omitted for clarity. B. The interactions of the molecular splint. The beta sheet in the little finger domain that is responsible for stabilizing the template strand is shown in stick representation. The splinting interactions between the little finger domain and template strand are shown as black dashes.

6.5. The templating base is displaced by phenanthriplatin in the extension complex

In the +1 extension complex, the phenanthridine moiety extends toward the 5´ end of the template strand (Fig. 22). The adduct is present in two conformations that are related by 180° rotation about Pt–N_P and 24° rotation about Pt–N_G. The flexibility about the Pt–N_P bond is the same type observed in a disordered small molecule structure solved by our collaborators, the Lippard lab at MIT[91]. The two lesion conformations in the macromolecular structure are discernible, but occupy overlapping space (Fig. 24A). As in the case of the insertion complex, the orientation of the phenanthridine ligand in the major conformation corresponds to the structure that was observed during previous studies with small molecule models of the phenanthriplatin-dG complex [91]. The two conformations of the phenanthridine ligand distort the templating dC differently, resulting in two conformations for this nucleotide. In the major lesion conformation (80%) the phenanthridine moiety would clash with the templating base dC if it were in its normal position and, consequently, the templating base is shifted by a ~75° propeller twist. In this twisted orientation, the templating base forms π-π stacking interactions with the phenanthridine rings. The dA(+2) residue located on the template strand downstream of the lesion forms a base stacking interaction with the phenanthridine ligand opposite the templating dC. Distortion of the templating base prevents it from forming a Watson-Crick base pair with the incoming nucleotide and may thus explain the large increase in K_M of the +1 extension step in the presence of the adduct. In the minor conformation (20%),

the orientation of the phenanthridine moiety permits the templating dC to assume a near-normal position, with only a slight twist to form a planar Watson-Crick base pair with the incoming nucleotide (Fig. 24B).

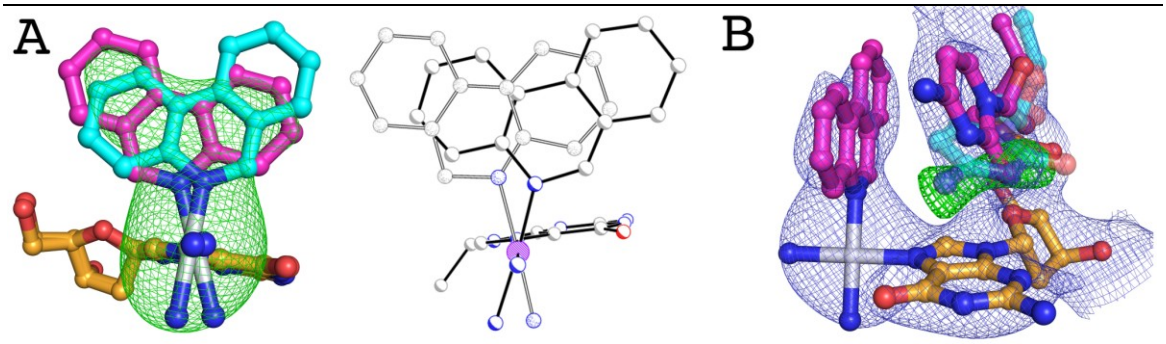


Figure 24. Double conformations of the phenanthriplatin adduct and templating base in the +1 extension complex

A. Two conformations of phenanthriplatin in the +1 extension complex and *cis*-$[Pt(NH_3)_2(Gua\text{-}Et)(Am)](OTf)_2$, where Gua-Et is a 9-ethylguanine, Am is phenanthridine, and OTf is trifluoromethanesulfonate. In the macromolecular structure (left), the major conformation is in magenta and minor conformation in cyan. The F_o-F_c omit map was calculated by removing the phenanthriplatin adduct and contoured at 3.0 σ. Two states of the disordered small molecule structure are shown in ball-and-stick mode (right). The small molecule structure was solved by collaborator Timothy Johnstone of the Steve Lippard lab at MIT. B. Template base minor conformation omit map. The F_o-F_c (green) and $2F_o$-F_c (blue) maps contoured at 3.0 σ and 1.0 σ, respectively. The omit map was calculated omitting the minor conformation, and the 2Fo-Fc map masks the structure of the major conformation (magenta). The minor conformation of the template dC is shown in semi-transparent cyan for reference.

6.6. Misalignment in the active site results in low efficiency of primer extension during the first extension step after phenanthriplatin-dG

Of the two alternate conformations, the minor species of the incoming dGMPNPP is shifted 0.5 Å into the minor groove (Fig. 25A). This shift is caused by hydrogen bond interaction with the shifted templating dC conformation (Fig. 25A). The major dGMPNPP conformation makes no hydrogen bonds with the templating dC, and is thus less shifted. The fidelity of Pol η in the +1 extension step suggests that the base pair interaction must be preserved, and therefore the minor conformation is probably the catalytically competent one (Fig. 18B).

In the major conformation, the platinated dG base is displaced ~1.3 Å into the major groove to avoid steric clash with the templating base. The primer strand terminal dC remains hydrogen bonded to the phenanthriplatin-dG, and through base pairing it pulls the primer 3´-OH to a distance 1.2 Å farther away from the active site than occurs in undamaged structures (Fig. 25B). The 3´-OH is no longer within coordination distance of the active site Mg^{2+}, nor is it close enough to the α-phosphate of the incoming nucleotide to participate in the phosphodiester bond formation (Fig. 25C) [27, 112]. Conversion from this catalytically incompetent (major) conformation to the competent (minor) conformation would require separation of the base pair between phenanthriplatin-dG and the primer. The energy required to break the base pair is probably a major contributor to the reduced k_{cat} of the +1 extension step.

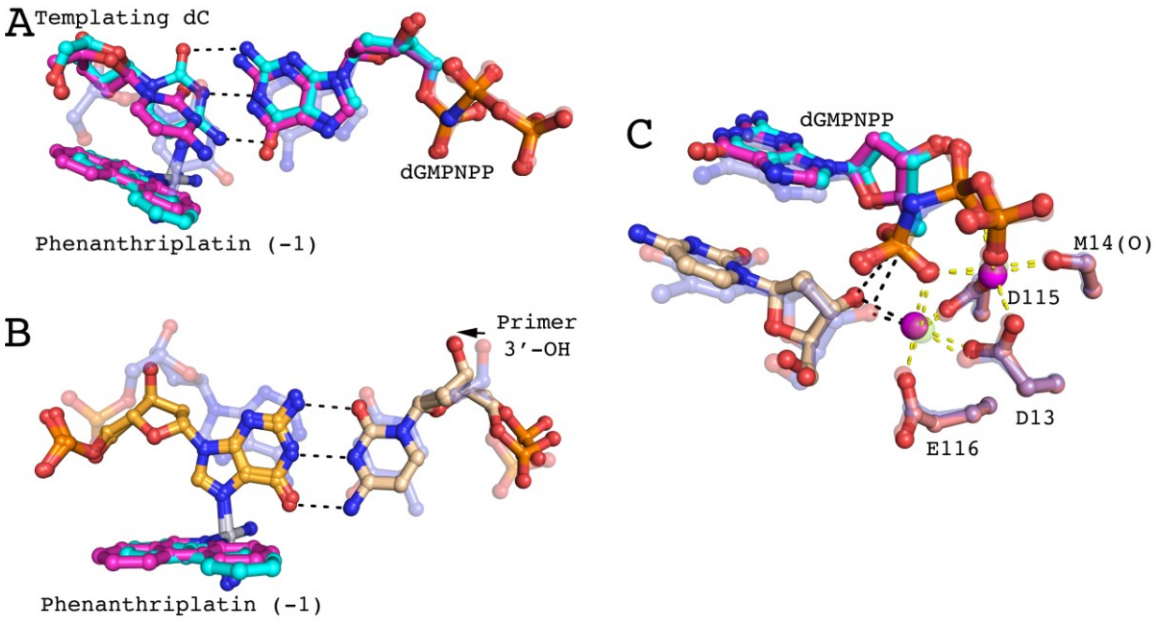


Figure 25. Misalignment of the active site in the +1 extension complex

A. Templating base pair of the +1 extension complex. In the minor conformation (cyan), hydrogen bonds are seen between the templating dC and incoming nucleotide (black dashes). The undamaged templating base pair is shown in semi-transparent blue for displacement reference. The phenanthriplatin adduct in the -1 position is also shown. B. Phenanthriplatin-dG:primer 3′ base pair of the +1 extension complex. The undamaged base pair is shown in semi-transparent blue for displacement reference. The 3′ base of the primer is shown in light orange and the adducted dG is shown in dark orange. The black dashes show the base pairing interaction and the arrow indicates the movement of the 3′-OH away from the active site. C. Primer misalignment in the extension step. Undamaged DNA-Pol η complex (3MR2) is shown in semi-transparent blue. The two conformations of the incoming nucleotide are shown in magenta and cyan

(major and minor conformations, respectively) and primer terminus in light orange. Yellow dashes indicate the coordination of the active site magnesium ions. The primer 3´-OH is displaced 0.7 Å from the undamaged position and the increased distances to the catalytic Mg^{2+} and α-phosphate of the incoming nucleotide are shown in black dashes.

6.7. DNA backbone displacement during phenanthriplatin TLS is accommodated by the CPD-binding pocket in the Pol η finger domain

In the +1 extension complex, the phenanthriplatin ligand occupies the site that is normally occupied by the phosphate backbone of the templating nucleotide in undamaged DNA-Pol η complexes (Fig. 26A). As a result, the phosphate backbone is moved 4.6 Å toward the fingers domain. This new backbone path is accommodated by the CPD-binding pocket above the template base and the enlarged active site, features not found in other polymerases.

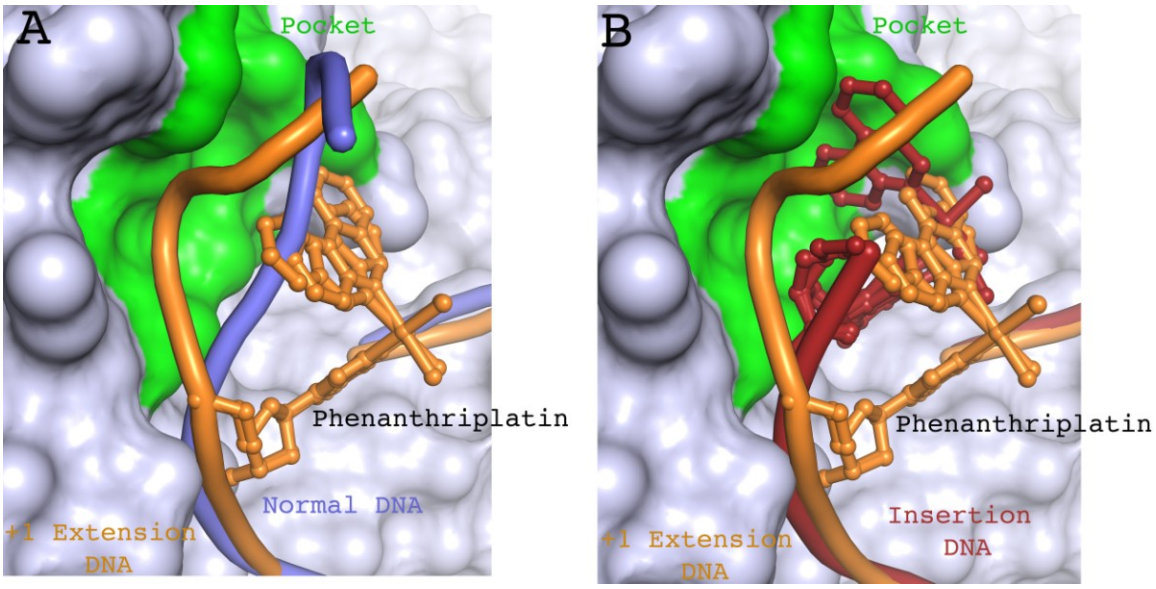


Figure 26. The CPD-binding pocket of the Pol η finger domain accommodates a large DNA backbone rearrangement in the +1 extension complex

A. Superposition of +1 extension complex with phenanthriplatin-damaged (orange) and undamaged DNA (light blue) bound to Pol η (4DL3). Cartoon traces the path of the DNA backbone and phenanthriplatin-dG is shown as sticks. Phenanthriplatin major and minor conformations are shown and both clash with the normal DNA backbone. The Pol η finger domain pocket is colored in green. B. Superposition of the phenanthriplatin-dG DNA from the +1 extension complex (orange) and the insertion complex DNA (red) bound to Pol η. Cartoon traces the path of the DNA backbone and phenanthriplatin-dG is shown as sticks. The Pol η finger domain pocket is colored in green. The pocket accommodates the

phenanthriplatin damage present in the insertion complex and the displaced DNA backbone of the +1 extension complex.

For example, the altered DNA backbone observed for Pol η, when modeled into the Pol κ structure, clashes with loop residues (aa. 133-135) of the Pol κ fingers domain (Fig. 27). In the +1 extension step, the expanded active site and finger domain pocket of Pol η are required in order to accommodate the platinum adduct, which may explain the strong stalling observed with all DNA polymerases other than Pol η (Fig. 17).

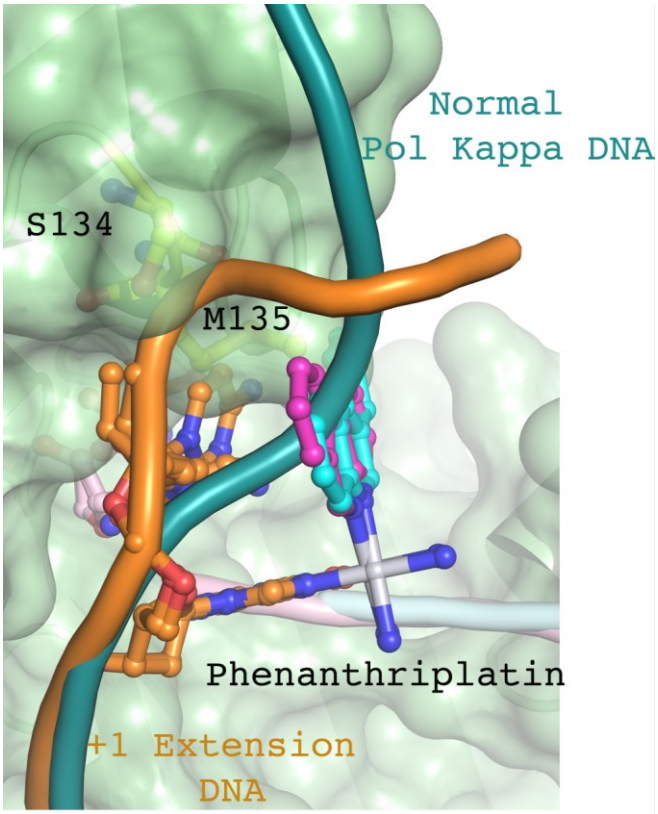


Figure 27. Model of the phenanthriplatin extension complex of Pol κ demonstrates steric clashes responsible for stalling

The Pol κ DNA is shown in teal and phenanthriplatin-damaged DNA from the Pol η extension complex is shown in orange. S134 (yellow) clashes with the backbone of the damaged DNA. M135 (yellow) clashes with the phenanthridine ligand and with the major conformation of the templating dC (orange sticks).

7. Summary and Conclusions

7.1. Active site alignment and structural transitions required for efficient polymerization of DNA

The first structure of Pol η was of the yeast Pol η polymerase, solved by the Aggrawal group[68]. This initial structure lacked DNA, metal ions, or incoming nucleotide but confirmed that despite low sequence conservation with polymerases of the A- and B-families, Pol η shared a similar fold[7, 68]. Six years later, a ternary yeast Pol η structure with DNA, metal (Ca^{2+} instead of Mg^{2+}), and incoming nucleotide was solved by the Carell group[113]. This structure, however, did not represent a catalytically active form, as evidenced by a modeled primer 3´-OH to α-phosphate distance of over 8Å. The first catalytically competent structure of Pol η was solved using human Pol η by the Yang group[25]. In the human Pol η structure, the active site resembled the architecture seen for polymerases of the A- and B-family polymerases[104, 114] with two Mg^{2+} ions coordinated by three conserved acidic residues (Asp13, Asp115, and Glu116). The 3´-OH in this structure is also coordinating a Mg^{2+} ion (A-site) and is only 3.3Å from the α-phosphate, consistent with distances observed in catalytically competent Pol β structures [29]. Time-resolved crystallographic structures of Pol η later demonstrated the primer 3´-OH only aligned for catalysis upon binding of the A-site Mg^{2+}[27]. Prior to Mg^{2+} binding, the 3´-OH was interacting with the S113 residue in a ground state (GS), but upon Mg^{2+} binding the 3´-OH transitioned to coordinating the A-site Mg^{2+} in a reactant state (RS), aligned for catalysis[27]. Structure based alignment of Y-family polymerases, however, reveal that S113 was a conserved position [102]. My research focus is on the alignment and

deprotonation of the 3´-OH in Pol η catalysis. Due to S113's proximity to the 3´-OH and conservation, I hypothesized that S113 may play a role in primer deprotonation or alignment[28].

In this work I investigated the role of the conserved S113 position within the Pol η active site. This site is also conserved across A- and B-family polymerases, and previous mutational studies of *E. coli* Pol I have identified reduced activity upon mutation, however, no function was identified[54, 55]. Mutation of Pol η S113 also reduces activity by ~4-fold, but my traditional static polymerase-substrate complex crystal structures of the mutant failed to identify a cause for reduced activity. Time resolved structural studies of the mutant reveal two defects in the S113A mutant: primer terminal base instability prior to Mg^{2+} binding and inefficient transition from the GS to the RS. Conformational flexibility of the primer end in the active is much more pronounced than previous structural work suggested[27]. Organization of the primer terminus prior to Mg^{2+} binding is dependent upon S113 in Pol η through direct hydrogen bond between the serine hydroxyl and the 3´-OH, stabilizing the primer in the GS. Upon activation of the WT polymerase by binding of Mg^{2+} to the A-site, the primer transitions from the GS to RS, an essential pre-catalytic step that is not detectable using static crystal structures alone[27]. Time resolved structures of S113A mutant visualized the inefficient GS to RS transition within the mutant active site, which is responsible the mutant's reduced activity. The GS to RS transition therefore is not automatic, but facilitated by the rearrangement of the S113 hydrogen bond network within the active site. To investigate the effect of GS to RS transition, primers

terminated with RNA bases were used which are biased for the C3´endo sugar pucker conformation, due to a steric clash of the 2´-OH group with downstream nucleotides, in this case the incoming dATP. Polymerization reactions using RNA terminated primers rescue the S113A mutant activity, and time resolved structures of S113A mutant extending RNA demonstrate the activity is rescued due to RNA's intrinsic sugar pucker bias favoring the RS conformation. This confirms the GS to RS transition defect is responsible for reducing S113A activity.

Time resolved crystallography has enabled visualization of catalytic events and mutational effects previously not accessible with traditional crystallographic techniques. Direct visualization removes the uncertainty in interpretation of kinetic phenomena and identifies the precise nature of kinetic events. Though not applicable to all systems, further development and use of this technique can facilitate greater understanding of the subtle conformational changes critical to enzyme-catalyzed reactions.

7.2. The polymerase active site promotes primer 3'-OH deprotonation, allowing the proton acceptor to be varied

Despite considerable effort in the polymerase field, a consensus on the identity of the general base that extracts the 3´-OH for polymerase catalysis has not been reached[48-51]. This is primarily due to the intricate complexity of two metal ion catalysis, in which mutational effects of candidate residues on deprotonation versus coordinating the active site metal ions cannot be unambiguously parsed. In lieu of mutational approaches, many QM/MM studies

have suggested a water molecule, and not a protein residue, as the most energetically favorable group to directly receive the proton[50, 51]. The recently discovered water that transiently interacts with the 3´-OH immediately prior to catalysis in Pol η was a promising candidate to accept the 3´-OH proton[25]. When I blocked the transient water site with a fluorine group, which is incapable of accepting a proton, Pol η remained active demonstrating that the transient water is not the sole, direct proton acceptor. Only one other water is close enough to the 3´-OH to directly accept the proton, the water coordinating the A-site Mg^{2+}. There is no straightforward means by which to block this water, so only *in silico* simulations will be able to shed light on a possible path for the proton to the metal water. The alternative proton acceptors observed in the three different primer configurations with Pol η (normal primer, 2´-OH and 2´-F terminated primers) indicate that the route for the 3´-OH proton to depart is variable. The impetus for deprotonation is being supplied by stable coordination of the 3´-OH to the Mg^{2+} ions[35, 53, 109], lowering the energetic requirement for the strength of the general base to a level that multiple groups can facilitate deprotonation.

7.3. Polymerase η as a target for anti-cancer chemotherapy

The platinum class of chemotherapeutics induces DNA damage in order to inhibit DNA replication[76, 115]. Inhibition of replication allows for potent toxicity in rapidly dividing cancer cells, while mitigating effects on somatic tissues which rarely replicate. In somatic cells, which are no longer actively dividing, DNA adducts are removed by housekeeping DNA repair mechanisms[74]. In actively

dividing cancer cell lines many of these DNA adduct will persist into S-phase, and TLS will be required to avoid replication fork collapse and apoptosis[74]. Inhibition of TLS polymerases would increase the toxicity of platinum agents in cancer cells that rely on TLS and give rise to greater efficacy of chemotherapy treatment.

I have identified an inhibitory small molecule binding site in the TLS polymerase η. I showed that dATP binds weakly at this site and is capable of inhibiting Pol η. I have not yet identified a high affinity inhibitor for this site, but using dATP as an initial template, I hope that rational design approaches will yield higher affinity inhibitors. Such inhibitors, if used in conjunction with chemotherapy treatments such as cisplatin or oxaliplatin, whose efficacy is correlated to Pol η activity, could dramatically improve treatment outcomes.

7.4. Inhibition of Polymerase η by monofunctional and bifunctional platinum chemotherapeutics

The mechanism of action of current platinum-based chemotherapeutics is to damage DNA and inhibit cellular processes such as DNA replication to induce cell death[76]. My collaborators in the Lippard lab at MIT developed a new platinum-based chemotherapeutic candidate phenanthriplatin[87]. Phenanthriplatin is a derivative of the most commonly used platinum chemotherapeutic, cisplatin, and I hypothesized that it may have a similar mechanism of action. I investigated first the efficiency with which polymerases from the A (Klenow fragment & Pol ν), B (Pol ζ), and Y (Pol κ & Pol η) families can synthesize DNA past

phenanthriplatin-DNA damage to determine if phenanthriplatin stalls replication polymerases. Only the TLS polymerase Pol η is capable of fully bypassing the phenanthriplatin lesion. All other polymerases that I tested are able to insert a nucleotide opposite the phenanthriplatin-dG (the insertion step) but stall immediately after the lesion (the extension step), confirming that phenanthriplatin, like cisplatin, does inhibit high fidelity replicative polymerases[32]. Replication past the lesion by Pol η is inefficient but appears to be error-free [82, 116]. Structural studies of Pol η at the insertion and +1 extension step reveal the uniquely enlarged active site of Pol η, which is evolved to TLS bypass UV lesions[25], accommodates the phenanthriplatin lesion. Perturbation of the nascent base pair, primer alignment in the active site, and the downstream DNA conformation by phenanthriplatin adducts explains the inability to bypass phenanthriplatin as efficiently as bifunctional platinum adducts[32, 87]. These results highlight that the cellular processing of phenanthriplatin is distinct from that of bifunctional platinum compounds such as cisplatin and oxaliplatin, in which one polymerase is typically employed for the insertion step and another for the extension steps[32, 84, 94].

Increased Pol η expression is a known resistance mechanism employed in cancer cells against treatment with the current bifunctional platinum compounds, reducing their curative potential[82, 116]. Due to phenanthriplatin's much more efficient inhibition of Pol η, increased Pol η expression in cancer cells should confer significantly less survival advantage, suggesting phenanthriplatin may be useful against cancers that have already developed resistance, and may also

combat the development of resistance that limits cisplatin and oxaliplatin efficacy.

8. Appendix 1: Motility of kinesin motor proteins

8.1. Cellular functions of motor proteins

Eukaryotic cells are very large, highly structured systems. In order to maintain cellular structure and provide efficient transport over the long distances within cells they have evolved active transporters[117]. These intracellular transporters are the motor proteins kinesin, myosin, and dynein. Motor proteins move along the filaments, made of either actin or microtubules (MT), that comprise the cytoskeleton. Structural and single molecule studies of these motor proteins suggest that all three families use a mechanism similar to walking to produce movement along their cytoskeletal tracks[118, 119].

A variety of cargos rely on active transport by motor proteins. Within the secretory pathway vesicles must be transported between the endoplasmic reticulum (ER), golgi cisterns, and plasma membrane. Generally, vesicle movement between the ER and golgi is provided by dynein, while transport from the golgi to endosomes and the cell surface is provided by kinesins[117]. Active transport, as opposed to diffusion, of vesicles in the secretory pathway is vital to ensure they reach the proper destination. Active transport of daughter chromatids during mitosis is essential to ensure equal distribution of genomic content to the two daughter cells[120]. Mitotic kinesins provide the force for separation of the chromatids during anaphase by sliding MTs of the mitotic spindles past one another[120]. In this work, I focused on understanding the molecular mechanism underpinning motility of the kinesin family of motor proteins.

8.2. Kinesin structure and mechanism

All three motor protein families are thought to use a similar walking type mechanism to produce movement, often called a hand-over-hand mechanism[118]. Kinesin is the smallest and most well characterized of the three motor protein families, and has thus been the focus of efforts to describe the walking mechanism. It is likely that activity and regulatory themes important to kinesin mechanism of motility will be applicable to all three motor families.

Though kinesin structure varies between subtypes, the general structure of kinesin family proteins consists of three domains: the motor domain, the stalk domain, and the tail domain (Fig. 28A). In most kinesins the motor domain is located at the N-terminus of the protein[121]. The motor domain is the defining feature of the kinesin family and is highly conserved between 14 kinesin subfamilies[122]. It is a globular structure that houses both the ATPase active site as well as a MT binding site, and thus is also referred to as the MT binding head. The final few residues of the motor domain are known as the neck linker. The neck linker is a flexible stretch of ~13 residues that connects the globular core of the motor domain to the stalk domain. The stalk is dominantly α-helical domain, and the site of dimerization via the formation of a coiled-coil with another kinesin[123]. Electron microscopy has shown that the stalk domain of kinesin adopts an elongated conformation spanning up to 70nm in some cases[124]. Finally, the tail domain is located on the other end of the stalk and is responsible for binding of cargo. Certain kinesin tail domains bind directly to their cargo molecules, while others require adapter proteins often called kinesin light chains.

Kinesins produce movement along the microtubule in a defined direction, usually plus-end directed. As a dimer, a kinesin alternates which head is bound to the MT in an asymmetric hand-over-hand fashion similar to human walking[125, 126]. Each step results in a net 8 nm movement, the distance between α and β tubulin heterodimers, that is tightly coupled to the hydrolysis of a single ATP molecule[127]. Whether a head is bound or unbound to the MT is determined by its nucleotide state, i.e. the identity of the nucleotide bound in its active site (ATP, ADP, or none). The head is bound when empty or with ATP and unbound with ADP[128]. Thus, as a motor domain completes a full ATP hydrolysis cycle it will alternate between MT bound and unbound, completing a full step. Key to its long distance transport is its ability to take many steps along a single microtubule track without detaching. This ability is termed processivity. Single molecule studies of kinesin motors have shown that kinesins have strong processivity, allowing them to travel several microns in a single run[129, 130]. Processivity requires communication between the two motor domains to ensure that only the rear head detaches while the leading head remains bound to the MT to prevent complete disassociation.

The current model for communication between the MT binding heads relies on the neck linker and dimerization domain that connect them[118, 131-133]. It is thought that tension is produced in the linker regions between the heads when both are bound to the MT. The tension would then cause a conformational change in the rear head that leads to detachment from the MT[134]. The rear head would then rebind ahead of its partner to complete a step. Communication

between the heads in this model is thus dependent on the linker and stalk domains that connect them as well as the MT lattice structure. Unfortunately, the current model does not fully explain some observed kinesin activities. For example, in the absence of MT's, kinesin has both a low and high affinity MT binding site that would not be predicted of a homodimer[135]. Kinesin is also able to coordinate its binding to tubulin heterodimers in a fashion identical to binding to MTs, demonstrating that the heads are able to communicate in the absence of the MT lattice[128].

An additional mechanism is needed to explain the communication between heads that occurs independent of tethering between the heads or the MT lattice. Crystal structures of several kinesin MT binding heads have been previously solved. In multiple structures of the mitotic kinesin Eg5 the heads appear to form a dimeric structure that likely would influence the enzymatic activity of their ATPase sites[136, 137]. Such a direct interaction between heads could potentially explain the outlying data and represent a novel step in the kinesin walking mechanism. The goals of this work were to detect and identify complexes formed between kinesin head domains and characterize their role in coordination of the heads during kinesin movement.

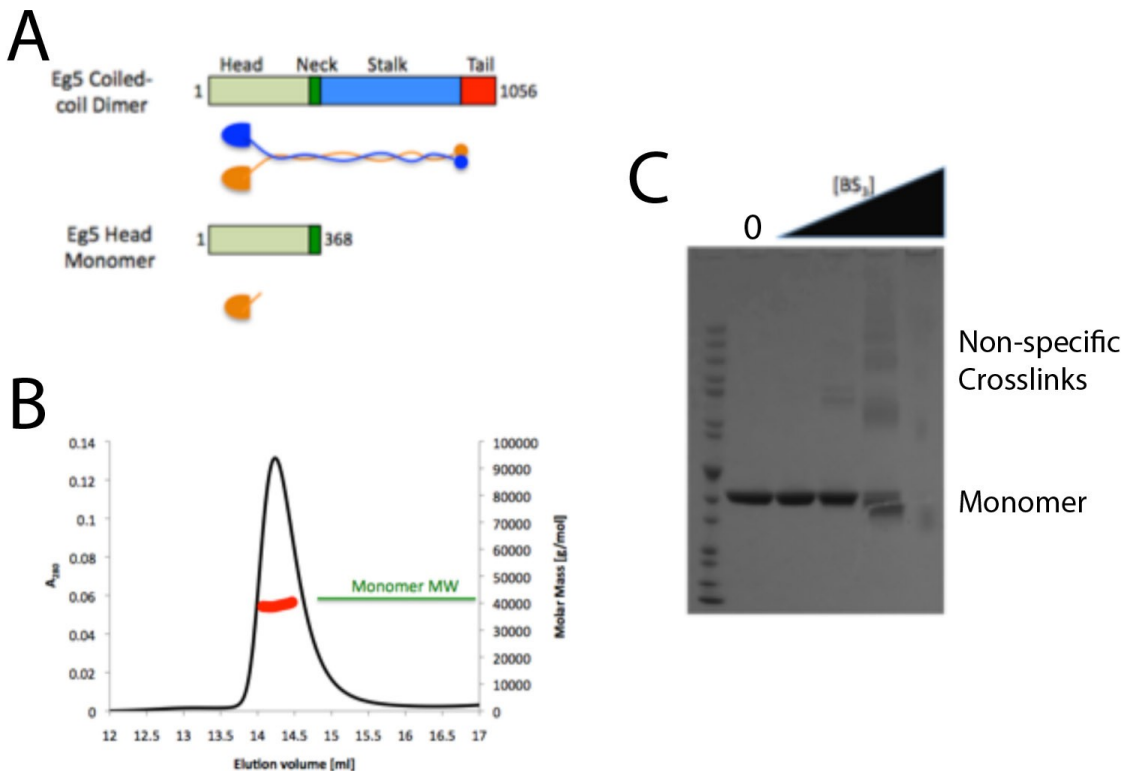


Figure 28. Kinesin structure and complex formation assays

A. Domains of WT and truncated Eg5. The general topology of the expressed protein is shown beneath. B. Gel filtration determined Eg5 constructs to be mono-disperse monomers in solution. Shown is a chromatograph of Eg5 monomers at ~8 mg/ml injection concentration and in the presence of ATP and ADP. All nucleotide conditions tested resulted in similar profiles. Multi-Angle Light Scattering (MALS) molecular weight measurement for the peak is shown in red and the predicted molecular weight of an Eg5 monomer is indicated in green. C. Crosslinking with BS3 cannot detect specific interaction between Eg5 monomers. Crosslinked bands are present at very high crosslinker

concentrations but form ladders, indicating these are non-specific crosslinks due to the highly concentrated reaction conditions.

8.3. Direct communication between kinesin motor domains is not detected by crosslinking and size exclusion chromatography

To enable detection of direct interactions between MT binding heads, a truncated construct of the mitotic kinesin Eg5 that lacks the coiled-coil dimerization domain and cargo binding domain was created (Fig. 28A). This construct retains only the MT binding head and has been shown to be monomeric in solution, which is essential for controlling its local concentration[138]. Importantly, the absence of the dimerization domain also eliminates any effects that tension may play in head communication. The construct is bacterially expressed and purified to homogeneity.

In order to detect dimerization interactions between the monomeric head constructs, size exclusion chromatography was used. Due to the nature of kinesin movement being linked to the ATPase cycles within the MT binding heads, any interaction would likely be dependent on the nucleotide state of the two heads involved. Size exclusion runs were performed with monomeric Eg5 heads with a variety of nucleotides (ATP, ADP, AMPNPP) as well as combinations of nucleotides to allow for detection of possible heterodimers. In all conditions, a single mono-disperse peak was observed (Fig. 28B). Multi-Angle Light Scattering (MALS) determine that these peaks corresponded to monomers in solution, indicating that no dimer formation was detectable.

Any interaction between kinesin heads during kinesin walking would need to be weak or transient to allow productive walking. Thus, I hypothesized that direct interaction between kinesin heads may be too transient or too weak to be detected by size exclusion chromatography. For increased sensitivity a protein crosslinking assay was used. Monomeric kinesin heads were crosslinked by BS3 in a variety of nucleotide conditions and in both the presence and absence of microtubules and depolymerized tubulin (Fig. 28C). Similarly no significant crosslinking was observed between kinesin heads in any condition tested.

8.4. Concentration dependent assays reveal no cooperativity between kinesin heads

Dimerization of kinesin heads to coordinate kinesin movements, no matter how transient or weak, would need to effect the ATPase activity of the heads in order to alter their MT binding state. In order to detect dimeric Eg5 head complexes that have the ability to alter the ATPase activity I measured the catalytic activity (k_{cat}) for ATP hydrolysis over a range of head domain concentrations. At higher concentrations dimerization would be promoted and catalytic effects due to complex formation would be observable. In initial experiments ATP hydrolysis appeared to increase sigmoidally with increasing head concentrations, consistent with a dimer complex forming at high concentrations that simulates the ATPase active site (Fig. 29A). Because the nucleotide state of the head determines the MT binding affinity of the head, such a complex could contribute to coordination of the heads during movement along

a microtubule. A similar phenomenon was observed for head domains of 4 other kinesins, KIF1A (monomeric family-3), DHC (dimeric family-1), KIF4B (dimeric family-4), and Ncd (dimeric family-14) (Fig. 29A). Upon addition of microtubules, kinesin activity is stimulated over 60 fold, but kinesin concentration dependent activation was still observed (Fig. 29B).

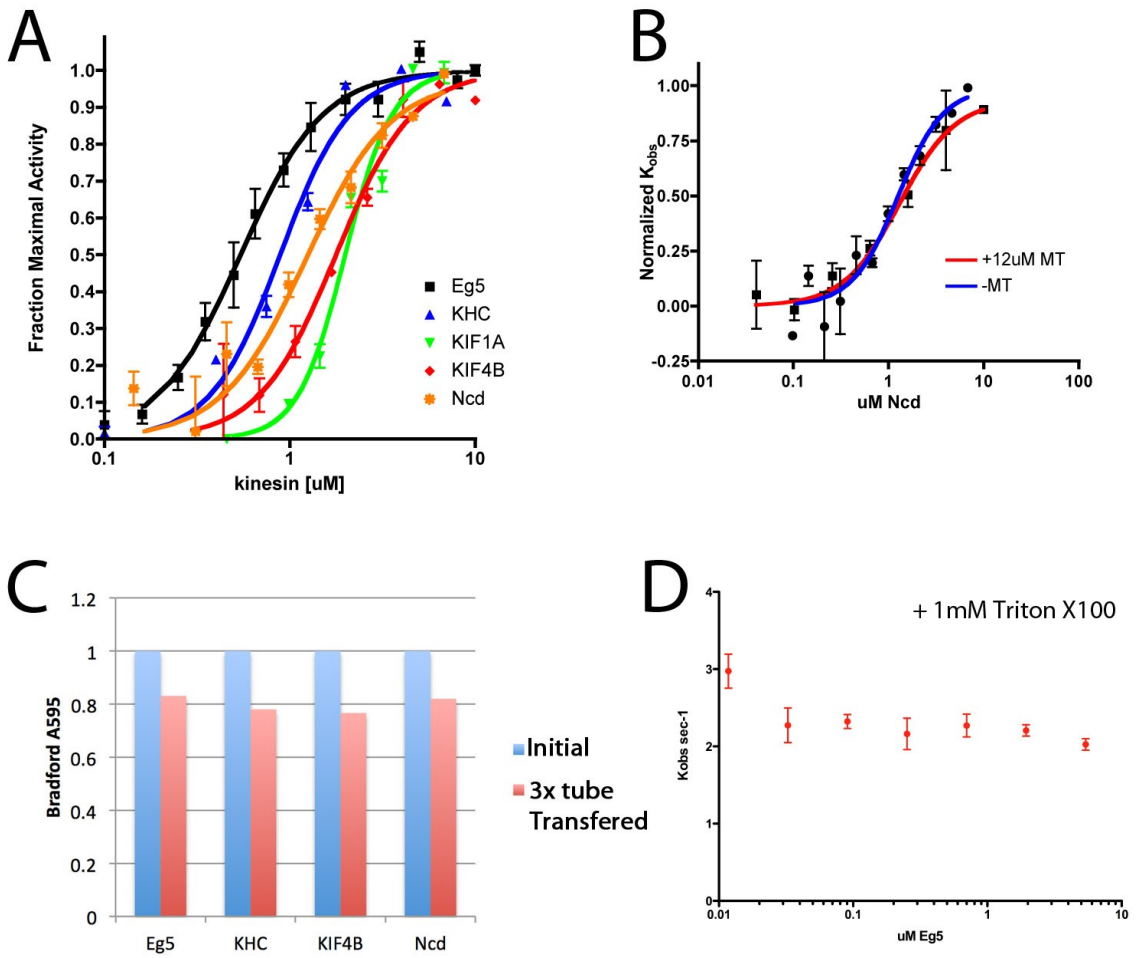


Figure 29. Concentration independent ATPase activity of kinesin monomers

A. Measurement of the steady state rate of ATP hydrolysis by kinesin monomers as a function of monomer concentration. A hill coefficient of ~2 or greater was determined for all curves by fitting the data with a sigmoidal dose response model, which is represented by the solid lines. B. Steady state ATP hydrolysis of

Ncd monomers in the presence and absence of polymerized microtubules (MT). Each curve is normalized independently due to the much higher overall rate in the presence of MTs. C. Protein losses due to tube adsorption detected by measuring protein concentration after transferring a solution of kinesin monomers through 3 tubes. All kinesin monomers had substantial losses due to adsorption despite ample (0.2 mg/ml) carrier BSA protein present in the buffer. D. Steady state ATP hydrolysis of Eg5 monomers in the presence of 1 mM Triton X100, a detergent that eliminates tube adsorption. Kinesin ATPase activity is independent of monomer concentration once tube adsorption is eliminated, indicating the concentration dependent activity observed previously was an artifact of the serial dilution protocol used.

Concentration dependent stimulation of kinesin activity appeared to demonstrate a regulatory kinesin head dimer could form, however in developing orthogonal assays (mutation based) this phenomenon was difficult to replicate and validate. For example, the degree of stimulation was dramatically reduced if samples of kinesin heads were prepared with parallel dilution methods instead of serial dilution. This lead me to be suspicious of dilution associated loss of protein concentration that could account for the apparent loss of activity at low concentrations, despite efforts to mitigate such losses with abundance BSA carrier protein in each reaction. To determine if kinesin concentration was lost during serial dilution, a sample of kinesin monomers was transferred between tubes three times and the concentration measured to access loss due to tube

binding (Fig. 29C). After only three transfers, Eg5 kinesin head concentration dropped dramatically, indicating the concentration dependent stimulation phenomenon observed was only a dilution artifact. This was confirmed by repeating the concentration dependence assay in the presence of 1 mM Triton X-100, a detergent I found that reduces tube adsorption of kinesin heads. In the presence of Triton, kinesin ATPase activity was independent of kinesin head concentration (Fig. 29D).

8.5. Summary

I originally hypothesized that direct interaction between the two microtubule binding heads of a kinesin moving along a microtubule would coordinate movement of the heads. Such coordination would require a direct interaction that affected the ATPase activity of the kinesin head, since the nucleotide state of a kinesin head is linked to it microtubule binding affinity. However, using size exclusion chromatography and protein crosslinking, no kinesin head dimerization interactions were detectable. Additionally, kinesin microtubule binding heads do not affect the ATPase activity of other heads, even at very high concentrations that would promote dimerization or interactions if they existed. These results suggest that a direct interaction between kinesin head domains does not form or contribute to coordination of kinesin movements. However, it is important to note that due to the nature of my negative results a direct interaction cannot be definitively ruled out.

8.6. Methods

Cloning and expression and purification of kinesin microtubule binding domains

The microtubule binding head of a representative set of kinesin molecules, HsEg5 (aa1-368), MmKIF1A (aa1-355), HsKIF4B (aa1-349), DmKHC (aa1-342), and DmNcd (aa335-700), were cloned and inserted into the pET28 or pET21 vector. The HsEg5 construct had no tag while all others had a C-terminal 6xhis tag. The HsEg5, MmKIF1A, and HsKIF4B constructs were expressed in BL21 (DE3) Rosetta cells by growing to an OD~0.8 at 37°C and inducing with 0.5 mM IPTG for 4 hrs. DmKHC and DmNcd constructs were expressed in BL21 (DE3) Rosetta cells by growing to an OD~0.8 at 37°C and inducing with 0.5 mM IPTG at 16°C for ~20 hr or 22°C for 8hrs, respectively. These kinesin domains are monomeric in solution.

Purification of HsEg5 began with lysing cells in lysis buffer (50 mM Na/Phosphate pH 6.8, 1 mM EDTA, 1 mM DTT, 1 mM AEBSF, 1 μM Leupeptin, Pepstatin, Benzamidine, 5 mM EDTA, 5 mM EGTA) by three or more passages through an Emulsiflex microfluidizer followed by clarification by ultracentrifugation at 35000rpm for 35 minutes. The lysate was loaded onto a Hitrap SP column equilibrated in 50 mM Phosphate pH 6.8, 1 mM EDTA, 1 mM DTT and eluted by 65ml gradient into 20 mM Phosphate pH 6.8, 1 M NaCl, 1 mM EDTA, 1 mM DTT. Fractions were pooled and diluted with 4 volumes of 50 mM Tris pH 8.5, 1 mM EDTA, 1 mM DTT buffer and loaded onto a MonoQ column. After ~3 column volumes of washing, fractions were eluted by gradient into 20 mM Tris pH 8.5, 1

M NaCl, 1 mM EDTA, 1 mM DTT. The pooled protein was then run on a superdex75 column in 20 mM PIPES pH 6.8, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT. The resulting protein was >98% pure.

Purification of DmKHC began with lysing cells in lysis buffer (25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5% glycerol, 20 mM imidazole, 1 mM AEBSF, 1 μM Leupeptin, Pepstatin, Benzamidine, 5 mM EDTA, 5 mM EGTA) by three or more passages through an Emulsiflex microfluidizer followed by clarification by ultracentrifugation at 35000rpm for 35 minutes. The lysate was loaded onto a His-Trap column equilibrated in 25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5% glycerol, 20 mM imidazole washed with 60 mM imidazole until UV absorbance stabilized and eluted in 300 mM imidazole. Fractions were pooled and exchanged into 25 mM Tris pH 8, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 2.5% glycerol buffer and loaded onto a MonoQ column. After ~3 column volumes of washing, fractions were eluted by 120ml gradient into 25 mM Tris pH 8, 1 M NaCl, 1 mM EDTA, 1 mM DTT, 2.5% glycerol. The pooled protein was then run on a superdex75 column in 20 mM PIPES pH 6.8, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT. The resulting protein was >98% pure.

Purification of HsKIF4B began with lysing cells in lysis buffer (25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5% glycerol, 20 mM imidazole, 1 mM AEBSF, 1 μM Leupeptin, Pepstatin, Benzamidine, 5 mM EDTA, 5 mM EGTA) by three or more passages through an Emulsiflex microfluidizer followed by clarification by ultracentrifugation at 35000rpm for 35 minutes. The lysate was loaded onto a His-Trap column equilibrated in 25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5%

glycerol, 20 mM imidazole washed with 60 mM imidazole until UV absorbance stabilized and eluted in 300 mM imidazole. Fractions were pooled and diluted with 9 volumes 25 mM PIPES pH 6.8, 1 mM EDTA, 1 mM DTT buffer and loaded onto a MonoS column equilibrated in the same buffer plus 50 mM NaCl. After ~3 column volumes of washing, fractions were eluted by 120ml gradient into 50 mM PIPES pH 6.8, 1 M NaCl, 1 mM EDTA, 1 mM DTT. The pooled protein was then run on a superdex75 column in 20 mM PIPES pH 6.8, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT. The resulting protein was >98% pure.

Purification of MmKIF1A began with lysing cells in lysis buffer (25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5% glycerol, 20 mM imidazole, 1 mM AEBSF, 1 μM Leupeptin, Pepstatin, Benzamidine, 5 mM EDTA, 5 mM EGTA) by three or more passages through an Emulsiflex microfluidizer followed by clarification by ultracentrifugation at 35000rpm for 35 minutes. The lysate was loaded onto a His-Trap column equilibrated in 25 mM Tris pH 8, 0.5M NaCl, 5 mM BME, 5% glycerol, 20 mM imidazole washed with 60 mM imidazole until UV absorbance stabilized and eluted in 300 mM imidazole. Fractions were pooled and diluted with 9 volumes 25 mM Hepes pH 7.5, 1 mM EDTA, 1 mM DTT buffer and loaded onto a MonoS column equilibrated in the same buffer plus 175 mM NaCl. After ~3 column volumes of washing, fractions were eluted by 90ml gradient into 25 mM Hepes pH 7.5, 1 M NaCl, 1 mM EDTA, 1 mM DTT. The pooled protein was then run on a superdex75 column in 20 mM PIPES pH 6.8, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT. The resulting protein was >98% pure.

Purification of DmNcd began with lysing cells in lysis buffer (10 mM Phosphate pH 7.2, 0.1 M NaCl, 2 mM $MgCl_2$, 1 mM DTT, 1 mM AEBSF, 1 µM Leupeptin, Pepstatin, Benzamidine, 5 mM EDTA, 5 mM EGTA) by three or more passages through an Emulsiflex microfluidizer followed by clarification by ultracentrifugation at 35000rpm for 35 minutes. The lysate was loaded onto a Hitrap SP column equilibrated in 10 mM Phosphate pH 7.2, 0.1 M NaCl, 2 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT and eluted by gradient into 10 mM Phosphate pH 7.2, 1 M NaCl, 2 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT. Fractions were pooled, dialyzed against 10 mM Tris pH 7.5, 80 mM NaCl, 2 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT buffer and loaded onto a HiTrapQ column. After ~3 column volumes of washing, fractions were eluted by 60ml gradient into 10 mM Tris pH 7.5, 1 M NaCl, 2 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT. The pooled protein was then run on a superdex75 column in 20 mM PIPES pH 6.8, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT. The resulting protein was >98% pure.

Crosslinking of kinesin microtubule binding domains

~1 mg/ml DmKHC monomers were incubated for 20 minutes with 0.05-1 mM BS3 crosslinker, 0.5 mM nucleotide (either ADP, AMPPNP, ATP, or and ADP/AMPPNP mixture), and ~2.5 mM tubulin (polymerized or depolymerized) where indicated. The total reaction volume was 10ul. The reaction buffer consisted of 50 mM PIPES pH 6.8, 5 mM MgCl2, 1 mM EGTA, and 1 mM DTT. Crosslinking reactions were quenched by addition of 2 ul 0.25 M Tris pH 7.4 for 5 minutes and run on a 3-8% Tris-Acetate gel followed by visualization with

coomassie blue or western blotting against the C-terminal 6xhis tag of the KHC monomers.

Concentration dependent kinesin ATPase activity assay

Purified kinesin monomers were serially diluted over a range of 10 μM to 0.1 μM in reaction buffer (20 mM Pipes pH 6.8, 50 mM KCl, 1 mM DTT, 5 mM $MgCl_2$, 0.2 mg/ml BSA) in an 8ul reaction volume. Where indicated 12 μM microtubules were also included, either polymerized or depolymerized. The reactions were initialized by addition of 2 μl 25 mM ATP, trace labeled with radioactive α-^{32}P ATP. After a short incubation (time varies based on activity level) the reactions were quenched with 100 mM EDTA and run on a TLC plate with the running buffer of 0.5 LiCl and 1 M formic acid. Reactions were visualized by exposure to a storage phosphor screen and scanning on a Typhoon Trio (GE Healthcare), quantified using ImageQuantTL (GE Healthcare) software, and curve fitting was performed using Prism 5 (Graphpad) software. Reported errors are calculated standard error of the mean (SEM) from triplicate experiments. For experiments with apparent sigmoidal activity profiles the data was fit with the equation Y=Bmax*X^h/(Kd^h + X^h), where Y is the activity, X is the protein concentration, Bmax is the maximal activity, and h is the hill coefficient.

9. Appendix 2: X-ray data collection and refinement statistics

9.1. Data collection and refinement statistics

Table 4. Data collection and refinement statistics for static ternary complex structures of WT Pol η

	WT+dA	WT+rA	WT+2´-Fl-dA	WT+dT low nucleotide
Data Collection				
Space group	P 61	P 61	P 61	P 61
Cell dimensions				
a, b, c (Å)	98.65	98.44	98.44	98.59
	98.65	98.44	98.44	98.59
	81.99	82.11	82.11	82.20
Wavelength (Ã)	1.00	1.00	1.00	1.00
Resolution (Ã)	30.04 - 1.57	21.31 - 1.45	21.31 - 1.31	30.04 - 1.80
Rsym (%)*	6.9 (47.7)	5.9 (56.5)	5.7 (68.9)	9.5 (71.8)
$I/\sigma I$*	11.70 (2.85)	9.02 (1.74)	12.03 (1.99)	10.57 (1.70)
Completeness (%)	99.69 (98.64)	97.59 (84.29)	96.77 (90.42)	90.68 (93.13)
Redundancy*	4.2 (3.3)	2.9 (1.7)	3.9 (3.5)	4.1 (4.0)
Refinement				
No. reflections	62082	76521	104816	38173
R_{work}	17.6	17.9	18.3	17.8
R_{free}	21.5	21.5	21.1	23.0
No. atoms	4412	4429	4491	4388
macromolecules	3881	3851	3912	3927
ligands	55	61	61	61
water	476	517	517	400
Average B-factor	20.6	23.7	20.3	21.7
macromolecules	19.6	22.6	19.3	21
solvent	27.5	31.1	26.6	27.8
R.m.s deviations				
Bond lengths (Å)	0.008	0.007	0.007	0.008
Bond angles (°)	1.20	1.14	1.18	1.19

* Highest resolution shell is shown in parenthesis

1 mM incoming nucleotide was used unless noted as "low nucleotide", in which case 0.1 mM was used

Table 5. Data collection and refinement statistics for static ternary complex structures of S113A mutant Pol η

	S113A+dT	S113A+dT low nucleotide	S113A+dA	S113A+rA
Data Collection				
Space group	P 61	P 61	P 61	P 61
Cell dimensions				
a, b, c (Å)	98.93	98.68	98.77	98.65
	98.93	98.68	98.77	98.65
	81.81	81.83	81.84	81.74
Wavelength (Ã)	1.00	1.00	1.00	1.00
Resolution (Ã)	32.38 - 2.26	31.5 - 1.85	31.51 - 1.8	23.36 - 1.81
Rsym (%)*	15.4 (68.6)	8.2 (62.8)	12.5 (66.1)	9.1 (54.4)
*I/σI**	8.53 (1.98)	10.31 (2.01)	9.17 (2.52)	10.01 (1.97)
Completeness (%)	98.27 (99.39)	99.97 (99.97)	99.60 (98.03)	97.73 (95.32)
Redundancy*	3.9 (3.7)	3.7 (3.5)	3.8 (3.8)	3.1 (2.2)
Refinement				
No. reflections	20856	38710	41915	40295
R_{work}	15.4	17.8	19.6	16.7
R_{free}	24.4	23.0	25.2	21.6
No. atoms	4363	4416	4435	4385
macromolecules	3879	3972	3942	3908
ligands	79	73	79	43
water	405	371	414	434
Average B-factor	25.8	22.9	20.6	22.4
macromolecules	25	22.2	19.8	21.6
solvent	32.4	28.9	27.3	28.7
R.m.s deviations				
Bond lengths (Å)	0.008	0.008	0.008	0.008
Bond angles (°)	1.17	1.15	1.19	1.13

* Highest resolution shell is shown in parenthesis

1 mM incoming nucleotide was used unless noted as "low nucleotide", in which case 0.1 mM was used

Table 6. Data collection and refinement statistics for time-resolved S113A mutant Pol η structures extending a dT terminated primer

	S113A+dT 0s	S113A+dT 480s
Data Collection		
Space group	P 61	P 61
Cell dimensions		
a, b, c (Å)	98.63	98.22
	98.63	98.22
	82.02	81.73
Wavelength (Ã)	1.00	1.00
Resolution (Ã)	26.91 - 1.46	29.48 - 1.94
Rsym (%)*	6.2 (60.6)	8.0 (48.0)
*I/σI**	9.08 (2.04)	9.47 (3.33)
Completeness (%)	98.99 (97.07)	98.42 (94.33)
Redundancy*	3.3 (3.1)	3.3 (2.6)
Refinement		
No. reflections	77578	32770
R_{work}	17.9	16.2
R_{free}	21.3	22.4
No. atoms	4645	4651
macromolecules	4004	4013
ligands	49	51
water	592	587
Average B-factor	22.2	29.4
macromolecules	21.1	28.3
solvent	29.2	36.4
R.m.s deviations		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	1.15	1.08

* Highest resolution shell is shown in parenthesis

Incoming nucleotide concentration was 1mM for all timepoints

Table 7. Data collection and refinement statistics for time-resolved S113A mutant Pol η structures extending a dA terminated primer

	S113A+dA 0s	S113A+dA 40s	S113A+dA 80s	S113A+dA 140s	S113A+dA 300s
Data Collection					
Space group	P 61	P 61	P 61	P 61	P 61
Cell dimensions					
a, b, c (Å)	98.30	98.10	98.36	98.16	98.32
	98.30	98.10	98.36	98.16	98.32
	81.83	82.13	81.75	82.51	81.94
Wavelength (Ã)	1.00	1.00	1.00	1.00	1.00
Resolution (Ã)	85.27 - 1.55	29.93 - 1.47	25.97 - 1.75	29.94 - 1.55	29.97 - 1.66
Rsym (%)*	5.6 (58.2)	6.0 (67.6)	7.3 (77.1)	7.1 (68.1)	6.8 (65.9)
*I/σI**	10.33 (2.16)	10.96 (2.50)	10.36 (2.14)	11.33 (2.17)	10.19 (2.29)
Completeness (%)	96.47 (92.41)	99.85 (99.47)	95.88 (92.16)	99.82 (99.40)	99.24 (99.05)
Redundancy*	3.2 (2.6)	3.7 (3.5)	3.8 (3.3)	3.8 (3.7)	3.6 (3.4)
Refinement					
No. reflections	62857	76205	43378	65439	52529
R_{work}	17.1	16.9	16.7	17.1	17.0
R_{free}	20.9	20.2	21.9	20.8	21.3
No. atoms	4516	4601	4570	4523	4541
macromolecules	3985	4034	4038	3995	4014
ligands	49	51	51	61	61
water	481	516	481	466	465
Average B-factor	24.7	21.2	26.1	20.6	25.4
macromolecules	23.9	20.4	25.4	19.7	24.7
solvent	30.4	27.3	31.6	27.4	30.9
R.m.s deviations					
Bond lengths (Å)	0.007	0.007	0.008	0.007	0.007
Bond angles (°)	1.12	1.13	1.09	1.10	1.12

* Highest resolution shell is shown in parenthesis

Incoming nucleotide concentration was 1mM for all timepoints

Table 8. Data collection and refinement statistics for time-resolved S113A mutant Pol η structures extending an rA terminated primer

	S113A+rA 0s	S113A+rA 40s	S113A+rA 140s	S113A+rA 230s	S113A+rA 300s
Data Collection					
Space group	P 61	P 61	P 61	P 61	P 61
Cell dimensions					
a, b, c (Å)	99.14	98.39	98.71	98.69	98.42
	99.14	98.39	98.71	98.69	98.42
	82.01	82.03	81.68	81.66	81.62
Wavelength (Ã)	1.00	1.00	1.00	1.00	1.00
Resolution (Ã)	32.45 - 1.83	29.98 - 1.91	31.46 - 1.85	30.04 - 1.80	29.97 - 1.92
Rsym (%)*	7.9 (66.2)	7.7 (58.7)	7.4 (61.1)	6.8 (62.4)	8.4 (69.4)
$I/\sigma I$*	10.28 (2.07)	8.92 (1.69)	12.34 (2.58)	13.99 (2.17)	11.62 (2.17)
Completeness (%)	98.53 (98.26)	99.48 (99.14)	98.80 (97.25)	99.92 (99.50)	99.80 (99.94)
Redundancy*	3.2 (3.2)	2.7 (2.3)	4.1 (4.1)	3.8 (3.6)	3.7 (3.6)
Refinement					
No. reflections	39758	34876	38360	41862	34312
R_{work}	16.6	16.2	16.1	16.4	16.0
R_{free}	21.7	22.1	21.4	20.9	21.9
No. atoms	4450	4506	4499	4519	4504
macromolecules	3942	3969	3973	3977	3979
ligands	49	60	61	61	60
water	459	477	465	481	465
Average B-factor	26.9	26.9	29	27.2	28.9
macromolecules	26.2	26.3	28.4	26.5	28.3
solvent	31.9	31.3	33.8	32.3	33.2
R.m.s deviations					
Bond lengths (Å)	0.008	0.008	0.008	0.008	0.008
Bond angles (°)	1.11	1.12	1.07	1.06	1.06

* Highest resolution shell is shown in parenthesis
Incoming nucleotide concentration was 1mM for all timepoints

Table 9. Data collection and refinement statistics for time-resolved WT Pol η structures extending a dA terminated primer

	WT+dA 0s	WT+dA 40s	WT+dA 100s	WT+dA 230s	WT+dA 300s
Data Collection					
Space group	P 61	P 61	P 61	P 61	P 61
Cell dimensions					
a, b, c (Å)	98.75	98.78	98.84	99.15	98.68
	98.75	98.78	98.84	99.15	98.68
	82.16	82.06	82.04	81.76	82.16
Wavelength (Ã)	1.00	1.00	1.00	1.00	1.00
Resolution (Ã)	32.33 - 1.8	30.08 - 1.82	32.3 - 1.98	30.16 - 2.05	32.1 - 2.07
Rsym (%)*	13.1 (72.4)	11.8 (68.0)	7.9 (56.7)	9.0 (64.1)	12.6 (61.7)
*I/σI**	10.72 (2.01)	10.13 (3.20)	13.32 (2.57)	9.89 (2.05)	11.72 (2.32)
Completeness (%)	98.63 (100.00)	99.51 (98.36)	99.76 (100.00)	98.85 (99.26)	99.82 (100.00)
Redundancy*	3.0 (3.1)	5.4 (5.3)	4.4 (4.4)	3.8 (3.6)	3.8 (3.7)
Refinement					
No. reflections	41632	40320	31727	28502	27678
R_{work}	17.4	17.5	15.5	16.7	15.8
R_{free}	22.7	22.7	21.7	23.0	23.0
No. atoms	4420	4394	4383	4237	4373
macromolecules	3881	3817	3825	3823	3824
ligands	49	60	61	61	61
water	490	517	497	353	488
Average B-factor	25.8	23	29.3	37.2	26.7
macromolecules	25.2	22.1	28.6	36.8	26.2
solvent	30.4	28.5	34.1	40.8	30.1
R.m.s deviations					
Bond lengths (Å)	0.008	0.008	0.008	0.008	0.008
Bond angles (°)	1.10	1.09	1.04	1.11	1.06

* Highest resolution shell is shown in parenthesis

Incoming nucleotide concentration was 1mM for all timepoints

Table 10. Data collection and refinement statistics for Pol η in complex with phenanthriplatin damaged DNA

	PPG Insertion	PPG +1 extension
Data Collection		
Space group	P 6_1	P 6_1
Cell dimensions		
a, b, c (Å)	98.77	98.69
	98.77	98.69
	82.20	81.81
Wavelength (Ã)	1.00	1.00
Resolution (Ã)	30-1.55	30-2.80
Rsym (%)*	7.8 (66.0)	8.9 (55.9)
*I/σI**	12.98 (1.81)	12.58 (1.78)
Completeness (%)*	99.7 (98.6)	99.6 (97.52)
Redundancy*	3.4 (3.2)	3.5 (2.9)
Refinement		
No. reflections	65924	11236
R_{work}	18.2	19.6
R_{free}	22.6	24.3
No. atoms	4335	3888
macromolecules	3890	3755
ligands	48	94
water	397	39
Average B-factor	22.5	35.9
macromolecules	22.0	36.1
solvent	26.5	28.1
R.m.s deviations		
Bond lengths (Å)	0.010	0.003
Bond angles (°)	1.42	0.84

* Highest resolution shell is shown in parenthesis

10. **References**

10.1. Bibliography

1. Kornberg A, McElroy WD, & Glass B (1957) The Chemical Basis of Heredity. *Johns Hopkins University Press*:579–608.
2. Loeb LA & Monnat RJ (2008) DNA polymerases and human disease. *Nat Rev Genet* 9(8):594-604.
3. Kornberg A, Lehman IR, Bessman MJ, & Simms ES (1956) Enzymic Synthesis of Deoxyribonucleic Acid. *Biochim Biophys Acta* 21(1):197-198.
4. Watson JD & Crick FHC (1953) Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* 171(4356):737-738.
5. Lehman IR*, et al.* (1958) Enzymatic Synthesis of Deoxyribonucleic Acid .5. Chemical Composition of Enzymatically Synthesized Deoxyribonucleic Acid. *P Natl Acad Sci USA* 44(12):1191-1196.
6. Meselson M & Stahl FW (1958) The Replication of DNA in Escherichia-Coli. *P Natl Acad Sci USA* 44(7):671-682.
7. Rothwell PJ & Waksman G (2005) Structure and mechanism of DNA polymerases. *Adv Protein Chem* 71:401-+.
8. Lehman IR (2003) Discovery of DNA polymerase. *J Biol Chem* 278(37):34733-34738.
9. Keen BA, Jozwiakowski SK, Bailey LJ, Bianchi J, & Doherty AJ (2014) Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic acids research*.
10. Johansson E & Dixon N (2013) Replicative DNA Polymerases. *Csh Perspect Biol* 5(6).
11. Ramadan K, Shevelev I, & Hubscher U (2004) The DNA-polymerase-X family: controllers of DNA quality? *Nat Rev Mol Cell Bio* 5(12):1038-1043.
12. Sale JE, Lehmann AR, & Woodgate R (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. *Nat Rev Mol Cell Bio* 13(3):141-152.
13. Blackburn EH & Collins K (2011) Telomerase: An RNP Enzyme Synthesizes DNA. *Csh Perspect Biol* 3(5).
14. Kati WM, Johnson KA, Jerva LF, & Anderson KS (1992) Mechanism and Fidelity of Hiv Reverse-Transcriptase. *J Biol Chem* 267(36):25988-25997.
15. Lipps G, Weinzierl AO, von Scheven G, Buchen C, & Cramer P (2004) Structure of a bifunctional DNA primase-polymerase. *Nat Struct Mol Biol* 11(2):157-162.
16. Ito J & Braithwaite DK (1991) Compilation and Alignment of DNA-Polymerase Sequences. *Nucleic acids research* 19(15):4045-4057.
17. Cannistraro VJ & Taylor JS (2004) DNA-thumb interactions and processivity of T7 DNA polymerase in comparison to yeast polymerase eta. *J Biol Chem* 279(18):18288-18295.
18. Johnson KA (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Bba-Proteins Proteom* 1804(5):1041-1048.
19. Kaushik N, Pandey VN, & Modak MJ (1996) Significance of the O-helix residues of Escherichia coli DNA polymerase I in DNA synthesis: Dynamics of the dNTP binding pocket. *Biochemistry-Us* 35(22):7256-7266.

20. Kunkel TA (1988) Exonucleolytic Proofreading. *Cell* 53(6):837-840.
21. Kunkel TA & Soni A (1988) Exonucleolytic Proofreading Enhances the Fidelity of DNA-Synthesis by Chick-Embryo DNA Polymerase-Gamma. *J Biol Chem* 263(9):4450-4459.
22. Li Y, Korolev S, & Waksman G (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *Embo J* 17(24):7514-7525.
23. Xia SL, Wang JM, & Konigsberg WH (2013) DNA Mismatch Synthesis Complexes Provide Insights into Base Selectivity of a B Family DNA Polymerase. *J Am Chem Soc* 135(1):193-202.
24. Aller P, Ye Y, Wallace SS, Burrows CJ, & Doublie S (2010) Crystal Structure of a Replicative DNA Polymerase Bound to the Oxidized Guanine Lesion Guanidinohydantoin. *Biochemistry-Us* 49(11):2502-2509.
25. Biertümpfel C*, et al.* (2010) Structure and mechanism of human DNA polymerase η. *Nature* 465(7301):1044-U1102.
26. Ren JS*, et al.* (1995) High-Resolution Structures of Hiv-1 Rt from 4 Rt-Inhibitor Complexes. *Nat Struct Biol* 2(4):293-302.
27. Nakamura T, Zhao Y, Yamagata Y, Hua Y-J, & Yang W (2012) Watching DNA polymerase η make a phosphodiester bond. *Nature* 487(7406):196-201.
28. Steitz TA (1999) DNA polymerases: Structural diversity and common mechanisms. *J Biol Chem* 274(25):17395-17398.
29. Batra VK*, et al.* (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* 14(4):757-766.
30. Storer AC & Cornish-Bowden A (1976) Concentration of MgATP2- and other ions in solution. Calculation of the true concentrations of species present in mixtures of associating ions. *The Biochemical journal* 159(1):1-5.
31. Tamasi G, Berrettini F, Hursthouse MB, & Cini R (2010) Effect of Free Water Molecules on the Structure of Mg-ATP- Dipyridylamine and Overview on Selected Metal-Adenosine Triphosphate Structures in Model Compounds and in Enzymes. *The Open Crystallography Journal* (3):1-13.
32. Zhao Y*, et al.* (2012) Structural basis of human DNA polymerase η-mediated chemoresistance to cisplatin. *P Natl Acad Sci USA* 109(19):7269-7274.
33. Freudenthal BD, Beard WA, Shock DD, & Wilson SH (2013) Observing a DNA Polymerase Choose Right from Wrong. *Cell* 154(1):157-168.
34. Brody RS & Frey PA (1981) Unambiguous Determination of the Stereochemistry of Nucleotidyl Transfer Catalyzed by DNA-Polymerase-I from Escherichia-Coli. *Biochemistry-Us* 20(5):1245-1252.
35. Steitz TA & Steitz JA (1993) A General 2-Metal-Ion Mechanism for Catalytic Rna. *P Natl Acad Sci USA* 90(14):6498-6502.
36. Brautigam CA & Steitz TA (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Current opinion in structural biology* 8(1):54-63.

37. Fiala KA & Suo Z (2004) Mechanism of DNA polymerization catalyzed by Sulfolobus solfataricus P2 DNA polymerase IV. *Biochemistry-Us* 43(7):2116-2125.
38. Arndt JW*, et al.* (2001) Insight into the catalytic mechanism of DNA polymerase beta: structures of intermediate complexes. *Biochemistry-Us* 40(18):5368-5375.
39. Shah AM, Li SX, Anderson KS, & Sweasy JB (2001) Y265H mutator mutant of DNA polymerase beta. Proper teometric alignment is critical for fidelity. *J Biol Chem* 276(14):10824-10831.
40. Washington MT, Prakash L, & Prakash S (2001) Yeast DNA polymerase eta utilizes an induced-fit mechanism of nucleotide incorporation. *Cell* 107(7):917-927.
41. Mizrahi V, Henrie RN, Marlier JF, Johnson KA, & Benkovic SJ (1985) Rate-limiting steps in the DNA polymerase I reaction pathway. *Biochemistry-Us* 24(15):4010-4018.
42. Rothwell PJ, Mitaksov V, & Waksman G (2005) Motions of the fingers subdomain of klentaq1 are fast and not rate limiting: implications for the molecular basis of fidelity in DNA polymerases. *Molecular cell* 19(3):345-355.
43. Beard WA & Wilson SH (2003) Structural insights into the origins of DNA polymerase fidelity. *Structure* 11(5):489-496.
44. Zhang H, Cao W, Zakharova E, Konigsberg W, & De La Cruz EM (2007) Fluorescence of 2-aminopurine reveals rapid conformational changes in the RB69 DNA polymerase-primer/template complexes upon binding and incorporation of matched deoxynucleoside triphosphates. *Nucleic acids research* 35(18):6052-6062.
45. Alberts IL, Wang Y, & Schlick T (2007) DNA polymerase beta catalysis: Are different mechanisms possible? *J Am Chem Soc* 129(36):11100-11110.
46. Schlick T, Arora K, Beard WA, & Wilson SH (2012) Perspective: pre-chemistry conformational changes in DNA polymerase mechanisms. *Theoretical chemistry accounts* 131:1287.
47. Showalter AK & Tsai MD (2002) A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry-Us* 41(34):10571-10576.
48. Batra VK*, et al.* (2013) Amino Acid Substitution in the Active Site of DNA Polymerase beta Explains the Energy Barrier of the Nucleotidyl Transfer Reaction. *J Am Chem Soc* 135(21):8078-8088.
49. Castro C*, et al.* (2007) Two proton transfers in the transition state for nucleotidyl transfer catalyzed by RNA- and DNA-dependent RNA and DNA polyrnerases. *P Natl Acad Sci USA* 104(11):4267-4272.
50. Wang LH, Yu XY, Hu P, Broyde S, & Zhang YK (2007) A water-mediated and substrate-assisted catalytic mechanism for Sulfolobus solfataricus DNA polymerase IV. *J Am Chem Soc* 129(15):4731-4737.
51. Wang Y & Schlick T (2008) Quantum mechanics/molecular mechanics investigation of the chemical reaction in Dpo4 reveals water-dependent

pathways and requirements for active site reorganization. *J Am Chem Soc* 130(40):13240-13250.
52. Cowan JA (1998) Metal Activation of Enzymes in Nucleic Acid Biochemistry. *Chemical reviews* 98(3):1067-1088.
53. Cowan JA (2002) Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine* 15(3):225-235.
54. Astatke M, Grindley NDF, & Joyce CM (1998) How E-coli DNA polymerase I (Klenow fragment) distinguishes between deoxy- and dideoxynucleotides. *J Mol Biol* 278(1):147-165.
55. Polesky AH, Steitz TA, Grindley NDF, & Joyce CM (1990) Identification of Residues Critical for the Polymerase-Activity of the Klenow Fragment of DNA-Polymerase-I from Escherichia-Coli. *J Biol Chem* 265(24):14579-14591.
56. Masutani C*, et al.* (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. *Nature* 399(6737):700-704.
57. Lehmann AR*, et al.* (1975) Xeroderma Pigmentosum Cells with Normal Levels of Excision Repair Have a Defect in DNA-Synthesis after Uv-Irradiation. *P Natl Acad Sci USA* 72(1):219-223.
58. Yamada A, Masutani C, Iwai S, & Hanaoka F (2000) Complementation of defective translesion synthesis and UV light sensitivity in xeroderma pigmentosum variant cells by human and mouse DNA polymerase eta. *Nucleic acids research* 28(13):2473-2480.
59. Cleaver JE (1972) Xeroderma Pigmentosum - Variants with Normal DNA-Repair and Normal Sensitivity to Ultraviolet-Light. *J Invest Dermatol* 58(3):124-&.
60. Robbins JH, Kraemer KH, & Flaxman BA (1975) DNA-Repair in Tumor-Cells from Variant Form of Xeroderma Pigmentosum. *J Invest Dermatol* 64(3):150-155.
61. Yang W (2014) An overview of Y-family DNA polymerases and a case study of human DNA pol eta. *Biochemistry-Us*.
62. Johnson RE, Prakash S, & Prakash L (1999) Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Pol eta. *Science* 283(5404):1001-1004.
63. Watanabe K*, et al.* (2004) Rad18 guides pol eta to replication stalling sites through physical interaction and PCNA monoubiquitination. *Embo J* 23(19):3886-3896.
64. Kunkel TA, Pavlov YI, & Bebenek K (2003) Functions of human DNA polymerases eta, kappa and iota suggested by their properties, including fidelity with undamaged DNA templates. *DNA Repair* 2(2):135-149.
65. Lovett ST (2007) Polymerase switching in DNA replication. *Molecular cell* 27(4):523-526.
66. Masuda Y, Piao J, & Kamiya K (2010) DNA Replication-Coupled PCNA Mono-Ubiquitination and Polymerase Switching in a Human In Vitro System. *J Mol Biol* 396(3):487-500.

67. Kannouche PL, Wing J, & Lehmann AR (2004) Interaction of human DNA polymerase eta with monoubiquitinated PCNA: A possible mechanism for the polymerase switch in response to DNA damage. *Molecular cell* 14(4):491-500.
68. Trincao J*, et al.* (2001) Structure of the catalytic core of S-cerevisiae DNA polymerase eta: Implications for translesion DNA synthesis. *Molecular cell* 8(2):417-426.
69. Pata JD (2010) Structural diversity of the Y-family DNA polymerases. *Bba-Proteins Proteom* 1804(5):1124-1135.
70. Park H*, et al.* (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *P Natl Acad Sci USA* 99(25):15965-15970.
71. O'Dwyer PJ, Stevenson JP, & Johnson SW (1999) Clinical status of cisplatin, carboplatin, and other platinum-based antitumor drugs. *Cisplatin - Chemistry and Biochemistry of a Leading Anticancer Drug*, ed Lippert B (Verlag Helvetica Chimica Acta, Zürich, Switzerland), pp 31-69.
72. Rudd GN, Hartley JA, & Souhami RL (1995) Persistence of Cisplatin-Induced DNA Interstrand Cross-Linking in Peripheral-Blood Mononuclear-Cells from Elderly and Young Individuals. *Cancer Chemoth Pharm* 35(4):323-326.
73. Branzei D & Foiani M (2005) The DNA damage response during DNA replication. *Curr Opin Cell Biol* 17(6):568-575.
74. Wang D & Lippard SJ (2005) Cellular processing of platinum anticancer drugs. *Nat Rev Drug Discov* 4(4):307-320.
75. Brabec V & Kasparkova J (2005) Modifications of DNA by platinum complexes: Relation to resistance of tumors to platinum antitumor drugs. *Drug Resist Updat* 8(3):131-146.
76. Kelland L (2007) The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer* 7(8):573-584.
77. Martin LP, Hamilton TC, & Schilder RJ (2008) Platinum resistance: The role of DNA repair pathways. *Clin Cancer Res* 14(5):1291-1295.
78. Todd RC & Lippard SJ (2009) Inhibition of transcription by platinum antitumor compounds. *Metallomics* 1(4):280-291.
79. Mamenta EL*, et al.* (1994) Enhanced replicative bypass of platinum-DNA adducts in cisplatin-resistant human ovarian-carcinoma cell-lines. *Cancer Res* 54(13):3500-3505.
80. Yang W & Woodgate R (2007) What a difference a decade makes: Insights into translesion DNA synthesis. *P Natl Acad Sci USA* 104(40):15591-15598.
81. Broyde S, Wang L, Rechkoblit O, Geacintov NE, & Patel DJ (2008) Lesion processing: High-fidelity versus lesion-bypass DNA polymerases. *Trends Biochem Sci* 33(5):209-219.
82. Ceppi P*, et al.* (2009) Polymerase η mRNA expression predicts survival of non-small cell lung cancer patients treated with platinum-based chemotherapy. *Clin Cancer Res* 15(3):1039-1045.

83. Albertella MR, Green CM, Lehmann AR, & O'Connor MJ (2005) A role for polymerase η in the cellular tolerance to cisplatin-induced damage. *Cancer Res* 65(21):9799-9806.
84. Hicks JK*, et al.* (2010) Differential roles for DNA polymerases eta, zeta, and REV1 in lesion bypass of intrastrand versus interstrand DNA cross-links. *Mol Cell Biol* 30(5):1217-1230.
85. Lovejoy KS & Lippard SJ (2009) Non-traditional platinum compounds for improved accumulation, oral bioavailability, and tumor targeting. *Dalton Trans* (48):10651-10659.
86. Johnstone TC, Wilson JJ, & Lippard SJ (2013) Monofunctional and higher-valent platinum anticancer agents. *Inorg Chem* 52(21):12234–12249.
87. Park GY, Wilson JJ, Song Y, & Lippard SJ (2012) Phenanthriplatin, a monofunctional DNA-binding platinum anticancer drug candidate with unusual potency and cellular activity profile. *P Natl Acad Sci USA* 109(30):11987-11992.
88. Johnstone TC, Park GY, & Lippard SJ (2014) Understanding and improving platinum anticancer drugs – Phenanthriplatin. *Anticancer Research* 34(1):471-476.
89. Kellinger MW, Park GY, Chong J, Lippard SJ, & Wang D (2013) Effect of a monofunctional phenanthriplatin-DNA adduct on RNA polymerase II transcriptional fidelity and translesion synthesis. *J Am Chem Soc* 135(35):13054-13061.
90. Johnstone TC, Alexander SM, Lin W, & Lippard SJ (2014) Effects of monofunctional platinum agents on bacterial growth: A retrospective study. *J Am Chem Soc* 136(1):116-118.
91. Johnstone TC & Lippard SJ (2014) The Chiral Potential of Phenanthriplatin and Its Influence on Guanine Binding. *J Am Chem Soc* 136(5):2126–2134.
92. Wiedemeyer WR, Beach JA, & Karlan BY (2014) Reversing Platinum Resistance in High-Grade Serous Ovarian Carcinoma: Targeting BRCA and the Homologous Recombination System. *Frontiers in oncology* 4:34.
93. Hoelz A, Nairn AC, & Kuriyan J (2003) Crystal structure of a tetradecameric assembly of the association domain of Ca(2+)/calmodulin-dependent kinase II. *Molecular cell* 11(5):1241-1251.
94. Lee YS, Gregory MT, & Yang W (2014) Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Accepted PNAS*.
95. French S & Wilson K (1978) On the treatment of negative intensity observations. *Acta Crystallographica Section A* 34(4):517-525.
96. Kabsch W (2010) Xds. *Acta Crystallogr D* 66:125-132.
97. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276:307-326.
98. Winn MD*, et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* 67:235-242.
99. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66:213-221.

100. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D* 66:486-501.
101. Schrodinger L (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
102. Boudsocq F, Ling H, Yang W, & Woodgate R (2002) Structure-based interpretation of missense mutations in Y-family DNA polymerases and their implications for polymerase function and lesion bypass. *DNA Repair* 1(5):343-358.
103. Brieba LG*, et al.* (2004) Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *Embo J* 23(17):3452-3461.
104. Wang F & Yang W (2009) Structural Insight into Translesion Synthesis by DNA Pol II. *Cell* 139(7):1279-1289.
105. Sievers F*, et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7.
106. Sponer J, Leszczynski J, & Hobza P (1996) Nature of nucleic acid-base stacking: Nonempirical ab initio and empirical potential characterization of 10 stacked base dimers. Comparison of stacked and H-bonded base pairs. *J Phys Chem-Us* 100(13):5590-5596.
107. Dickerson RE*, et al.* (1982) The Anatomy of a-DNA, B-DNA, and Z-DNA. *Science* 216(4545):475-485.
108. McElroy WD & Glass B (1951) *Phosphorus metabolism, a Symposium on the Role of Phosphorus in the Metabolism of Plants and Animals* (Johns Hopkins Press, Baltimore,).
109. Sissi C & Palumbo M (2009) Effects of magnesium and related divalent metal ions in topoisomerase structure and function. *Nucleic acids research* 37(3):702-711.
110. McCulloch SD*, et al.* (2004) Preferential cis-syn thymine dimer bypass by DNA polymerase eta occurs with biased fidelity. *Nature* 428(6978):97-100.
111. Lee Y-S, Gregory MT, & Yang W (2014) Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *P Natl Acad Sci USA* 111(8):2954–2959.
112. Hsin K, Sheng Y, Harding MM, Taylor P, & Walkinshaw MD (2008) MESPEUS: a database of the geometry of metal sites in proteins. *J Appl Crystallogr* 41:963-968.
113. Alt A*, et al.* (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science* 318(5852):967-970.
114. Doublie S, Tabor S, Long AM, Richardson CC, & Ellenberger T (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 angstrom resolution. *Nature* 391(6664):251-258.
115. Maccio A & Madeddu C (2013) Cisplatin: an old drug with a newfound efficacy - from mechanisms of action to cytotoxicity. *Expert Opin Pharmaco* 14(13):1839-1857.

116. Zhou WD*, et al.* (2013) Expression of DNA Translesion Synthesis Polymerase eta in Head and Neck Squamous Cell Cancer Predicts Resistance to Gemcitabine and Cisplatin-Based Chemotherapy. *Plos One* 8(12).
117. Brownhill K, Wood L, & Allan V (2009) Molecular motors and the Golgi complex: Staying put and moving through. *Semin Cell Dev Biol* 20(7):784-792.
118. Gennerich A & Vale RD (2009) Walking the walk: how kinesin and dynein coordinate their steps. *Curr Opin Cell Biol* 21(1):59-67.
119. Tang Y*, et al.* (1999) Cellular motor protein KIF-4 associates with retroviral Gag. *J Virol* 73(12):10508-10513.
120. Valentine MT, Fordyce PM, & Block SM (2006) Eg5 steps it up! *Cell division* 1:31.
121. Hirokawa N (1998) Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* 279(5350):519-526.
122. Martinez NW, Xue XX, Berro RG, Kreitzer G, & Resh MD (2008) Kinesin KIF4 Regulates Intracellular Trafficking and Stability of the Human Immunodeficiency Virus Type 1 Gag Polyprotein. *J Virol* 82(20):9937-9950.
123. Chandra R, Salmon ED, Erickson HP, Lockhart A, & Endow SA (1993) Structural and Functional Domains of the Drosophila Ncd Microtubule Motor Protein. *J Biol Chem* 268(12):9005-9013.
124. Heald R (2000) Motor function in the mitotic spindle. *Cell* 102(4):399-402.
125. Hackney DD (1994) Evidence for Alternating Head Catalysis by Kinesin during Microtubule-Stimulated Atp Hydrolysis. *P Natl Acad Sci USA* 91(15):6865-6869.
126. Turner J*, et al.* (2001) Crystal structure of the mitotic spindle kinesin Eg5 reveals a novel conformation of the neck-linker. *J Biol Chem* 276(27):25496-25502.
127. Parke CL, Wojcik EJ, Kim S, & Worthylake DK (2010) ATP Hydrolysis in Eg5 Kinesin Involves a Catalytic Two-water Mechanism. *J Biol Chem* 285(8):5859-5867.
128. Alonso MC*, et al.* (2007) An ATP gate controls tubulin binding by the tethered head of kinesin-1. *Science* 316(5821):120-123.
129. Block SM, Goldstein LSB, & Schnapp BJ (1990) Bead Movement by Single Kinesin Molecules Studied with Optical Tweezers. *Nature* 348(6299):348-352.
130. Hirokawa N & Noda Y (2008) Intracellular transport and kinesin superfamily proteins, KIFs: Structure, function, and dynamics. *Physiol Rev* 88(3):1089-1118.
131. Endow SA & Barker DS (2003) Processive and nonprocessive models of kinesin movement. *Annu Rev Physiol* 65:161-175.
132. Howard J, Hudspeth AJ, & Vale RD (1989) Movement of Microtubules by Single Kinesin Molecules. *Nature* 342(6246):154-158.
133. Naghavi MH & Goff SP (2007) Retroviral proteins that interact with the host cell cytoskeleton. *Curr Opin Immunol* 19(4):402-407.

134. McDonald D*, et al.* (2002) Visualization of the intracellular behavior of HIV in living cells. *J Cell Biol* 159(3):441-452.
135. Krauss J, de Quinto SL, Nusslein-Volhard C, & Ephrussi A (2009) Myosin-V Regulates oskar mRNA Localization in the Drosophila Oocyte. *Curr Biol* 19(12):1058-1063.
136. Mori T, Vale RD, & Tomishige M (2007) How kinesin waits between steps. *Nature* 450(7170):750-U715.
137. Schnitzer MJ & Block SM (1997) Kinesin hydrolyses one ATP per 8-nm step. *Nature* 388(6640):386-390.
138. Kim WK*, et al.* (1998) Binding of murine leukemia virus Gag polyproteins to KIF4, a microtubule-based motor protein. *J Virol* 72(8):6898-6901.

11. **Curriculum Vitae**

CURRICULUM VITAE

Mark T. Gregory

Laboratory of Molecular Biology
National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)
National Institutes of Health (NIH)
9000 Rockville Pike, Bldg 5, Rm B1-03, Bethesda, MD 20892
Phone: (301) 402-4647 Email: gregorymark@nih.gov

EDUCATION

	Degree	Year	Scientific field
Johns Hopkins University & National Institutes of Health	Ph.D.	2014	Biology, structural biology
University of Arizona	M.S.	2008	Biochemistry and biophysics
University of Arizona	B.S.	2007	Biochemistry and biophysics

RESEARCH EXPERIENCE

2008-present	Pre-Doctoral fellow in the laboratory of Dr. Wei Yang Laboratory of Molecular Biology, NIDDK, NIH Description: Structure and function of human polymerase η
2007-2008	Graduate researcher in the laboratory of Dr. Nancy Horton, Department of Chemistry and Biochemistry, University of Arizona Description: Biochemistry of archeal nucleotide excision repair (NER) proteins
2007	Summer Intern in the laboratory of Dr. Wei Yang Laboratory of Molecular Biology, NIDDK, NIH Description: Expression, purification, and structural studies of polymerase η
2006-2007	Undergraduate researcher in the laboratory of Dr. Nancy Horton Department of Chemistry and Biochemistry, University of Arizona Description: Structure of nucleotide excision repair (NER) protein XPC

SKILLS

Cloning, PCR, western/southern blotting, protein expression, immunoprecipitation and crosslinking assays, chromatography, x-ray crystallography

Mammalian tissue culture, sterile technique, cell transfection, basic Fluorescence-Activated Cell Sorting (FACS)

Experience with Unix/Linux systems, bash shell scripting, and a variety of structural data processing software (Phenix, Coot, Refmac, HKL2000, etc)

FELLOWSHIPS

2009-2011	Intramural AIDS Research Fellowship Office of AIDS Research, NIH

PUBLICATIONS

Gregory MT, Park GY, Lee Y, Johnstone TC, Yang W, Lippard SJ. Structural and mechanistic studies of polymerase η bypass of phenanthriplatin DNA damage. *Submitted PNAS.*

Gregory MT, Yang W. Mechanism of primer deprotonation and substrate alignment in DNA polymerase η catalysis. *In preparation.*

Lee YS, Gregory MT, Yang W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc Natl Acad Sci U S A. 2014 Jan 21.

Zhao Y, Gregory MT, Biertümpfel C, Hua YJ, Hanaoka F, Yang W. Mechanism of somatic hypermutation at the WA motif by human DNA polymerase η. Proc Natl Acad Sci U S A. 2013 May 14;110(20):8146-8151.

Zhao Y, Biertümpfel C, Gregory MT, Hua YJ, Hanaoka F, Yang W. Structural basis of human DNA polymerase η-mediated chemoresistance to cisplatin. Proc Natl Acad Sci U S A. 2012 May 8;109(19):7269-74.

Biertümpfel C, Zhao Y, Kondo Y, Ramón-Maiques S, Gregory M, Lee JY, Masutani C, Lehmann AR, Hanaoka F, Yang W. Structure and mechanism of human DNA polymerase eta. Nature. 2010 Jun 24;465(7301):1044-8.

Dunten PW, Little EJ, Gregory MT, Manohar VM, Dalton M, Hough D, Bitinaite J, Horton NC. The structure of SgrAI bound to DNA; recognition of an 8 base pair target. Nucleic Acids Res. 2008 Sep;36(16):5405-16.

TEACHING/MENTORING

2012 & 2013	Supervised summer interns at the National Institutes of Health
2008-2009	Teaching assistant at Johns Hopkins University Courses: Cell biology, Biochemistry
2007-2008	Supervised an undergraduate researcher at University of Arizona
2007-2008	Teaching assistant at University of Arizona Courses: Introductory biology, Biochemistry lab techniques
2004	Algebra and chemistry tutor at Brigham Young University

POSTER PRESENTATION

Gregory MT, Yang W. 2013. Time-resolved crystallography of polymerase catalysis. Keystone Symposium: DNA Replication and Recombination. Banff, Alberta, Canada.

Gregory MT, Yang W. 2010. Host motor proteins are involved in HIV intracellular motility. NIDDK fellow's retreat. Cambridge, MD.

PERSONAL ACTIVITIES

2010-2013	Founder and captain of a co-ed indoor soccer team
2008-present	Volunteer at the Montgomery County Manna Food Center